Identification and functional analysis of LCI15, a suppressor of the *air dier* phenotype of *LCIB* mutants in *Chlamydomonas reinhardtii*

by

Soujanya Akella

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Genetics

Program of Study Committee:
Martin H. Spalding, Co-major Professor
Steven R. Rodermel, Co-major Professor
David J. Oliver
Gustavo MacIntosh
Yanhai Yin

The student author, whose presentation of the scholarship herein was approved by the program of study committee, is solely responsible for the content of this dissertation. The Graduate College will ensure this dissertation is globally accessible and will not permit alterations after a degree is conferred.

Iowa State University

Ames, Iowa

2018

DEDICATION

To my parents,
for all their hard work, foresight, and sacrifice

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS

My journey towards this degree, while filled with many important people, has been made possible by my major professor, Dr. Martin Spalding, his kindness, and the faith he has had in me from the very beginning. Having arrived at Iowa State with a background in engineering (and next to none in biology), I had started out wanting some practical experience in a molecular biology lab. Dr. Spalding, despite likely having known that I had never even used a pipette before – much less ever formulated what goes into, say, a PCR (even in theory) – let me work in his lab so I could learn from scratch the basics of wet-lab work, when most others in his position would have (understandably) hesitated. He then gave me an even bigger opportunity – to work towards a graduate degree as his student – that has led to this experience of intellectual and personal growth and learning.

Over the years, Dr. Spalding has afforded me immense patience and understanding as I learnt to navigate, at my own pace, the worlds of Chlamy, the CCM, and of biology in general. His guidance, while extensive and regular, left room for flexibility that taught me to be independent. I am grateful for the confidence he has had and the potential he has seen in me this entire time. From all my years of learning from him and interacting with him, I have seen not just his passion for and expertise in what he does but also his pleasant disposition and forbearing personality. I aspire to one day be the kind of leader he is.

I am grateful to Dr. Steven Rodermel, Dr. David Oliver, Dr. Yanhai Yin, and Dr. Gustavo MacIntosh for being on my committee and for their helpful suggestions and incisive questions that added value to my research and made me a better scientist. Dr.

Rodermel – whose cell biology class that I took as part of my foundation in biology further piqued my interest in the subject – also helped me sort through my results particularly in the last couple of years of my lab work.

I'd also like to thank our collaborators – Dr. Sabeeha Merchant, Dr. Matteo Pellegrini, and Dr. Sorel Fitz-Gibbon at University of California, Los Angeles; and Dr. Donald Weeks and Dr. Wen Zhi Jiang at University of Nebraska–Lincoln – for their help and their contributions to my project. Sorel, who performed the sequencing and mapping work, has helped me understand the process and the results by patiently responding to my incessant questioning with many explanations and clarifications.

I will miss the company of everyone that I have known and worked alongside in our lab: Yingjun Wang, David Wright, Dan Stessman, Han Gao, Wei Fang, Bo Chen, and Alfredo Kono. From contemplating the meaning of life with Han on Friday evenings in the lab; getting reassured by Yingjun that my failed experiments would eventually work; joking with Wei about random stuff (sometimes ending up with both of us laughing *at me*, rather than us just laughing); accosting Bo a million times, some experiment protocol in hand, starting off a string of questions with "Bo, I have a question for you…"; having illuminating conversations about the current and past state of this country with David and Dan; to having discussions with Alfredo that start with us analyzing his results and end with me learning something about photosynthetic oxygen evolution and/or some protein (not to mention all the random chats about anything from contrasting cultures in our respective countries to the human gut microbiome and the effect of consumed sugar on it), there are countless memories that I will cherish and carry with me forever. I got so lucky to be amongst such genuinely friendly and well-meaning people every day.

A special mention goes out to Deqiang Duanmu, whom I knew for only my first 2-3 months in the lab when I followed him around on the fourth floor of Bessey Hall. He introduced me to benchwork, patiently teaching me how to perform genetic crosses of Chlamy and to extract genomic DNA (among more mundane lab procedures), while he himself was probably under a lot of stress being in the last stages of his experiments and working on his dissertation. I also have him to thank for generating the suppressor mutation in *pmp-su1*, the strain that came to define my years in our lab.

It was also a pleasure knowing Xiaowen Fei, Haili Dong, and Bo Xie, albeit for the relatively shorter periods of time that they were in our lab. Jane Roche and Boris Eyheraguibel, who took me under their wing when I was learning to drive, diligently gave me practice lessons (so naturally any praise – or criticism – for my driving skills should be directed towards them) during which we had many lively discussions about biology, research, and French culture.

My teaching life at Iowa State led me to Dr. Barbara Krumhardt, Mary Madsen, and Claudia Lemper, who did not just teach me about genetics, human anatomy and physiology, and the skill of pedagogy itself, but also enriched my time here with their friendship. All my years as a teaching assistant introduced me to students (from whom I myself learnt a lot) and other teaching assistants who would go on to become friends.

My parents, forever supportive and proud of me, have always been a source of enduring strength and compassion. This privileged life I am fortunate to lead is a result of their wisdom, hard work, and their devotion to their children. I owe all that I am, and all that I have been able to achieve, to them. The rest of my family in Iowa, elsewhere in the US, and in India, have encouraged me to work hard and do better. With my uncle and

aunt being in Ames, I have had access to comfort and support from family close by – something that I realize I'm lucky to have as an international student and for which I am very grateful to them.

My brother, Achyuta Akella – an unceasing source of support, inspiration, advice, laughs, and good company – has just always been there for me, sharing my perpetual love for (our late) dogs and putting up with all my tennis-related chatter. For the record – and no matter how many times he tries to claim otherwise – neither his French language skills nor his renditions of *my* pasta recipes are better than mine (and, no, he does not now suddenly "deserve" an honorary PhD). I would be remiss not to also acknowledge the inspiration and entertainment that I have gained from Roger Federer in particular and the sport of tennis in general over all these years.

There are many others in the Iowa State and Ames community – in the GDCB Department, Linda Wild in the Genetics program, in Bessey Hall, the on-campus Caribou Coffee – who have helped me at various times and have come to be my friends and acquaintances, making everyday life here more fun and interesting.

While I have experienced tough times as I worked toward this end, I have matured in the process and have seen successes sprout from persisting through those times. The long and winding road towards this PhD has taught me – perhaps far more than I realize in this moment – about the importance of failure, the power of perseverance, and the rewards ushered in by enduring hard work.

NOMENCLATURE

amiRNA	artificial microRNA
bp(s)	base pair(s)
C4	four carbon
CA or CAH	carbonic anhydrase
CC-####	culture collection number, as designated by the Chlamydomonas Resource Center (http://www.chlamycollection.org)
CCM	CO_2-concentrating mechanism
cDNA	complementary DNA
Ci	inorganic carbon
CO_2	carbon dioxide
COX2B	cytochrome c oxidase subunit 2B
dsDNA	double-stranded DNA
gDNA	genomic DNA
HCO_2	high CO_2
HCO_3^-	bicarbonate
LCI*X*	low CO_2-inducible *X*
LCO_2	low CO_2
mt^+/ mt^-	mating type *plus / minus*
O_2	oxygen
Par^R	paromomycin-resistant (or resistance)
PCR	polymerase chain reaction
ppm	parts per million

PRL1	pleiotrophic regulatory locus 1
RbcL	large subunit of Rubisco
Rubisco	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-1,5-bisphosphate
SNP	single nucleotide polymorphism
TAP	tris-acetate-phosphate (growth medium)
$VLCO_2$	very low CO_2
WT	wild type
Zeo^R	zeocin-resistant (or resistance)

ABSTRACT

The eukaryotic alga, *Chlamydomonas reinhardtii*, acclimates to limiting CO_2 conditions by the induction of the CO_2-concentrating mechanism (CCM) – a complex system of changes in its metabolism, gene expression patterns, and physiology – to compensate for the reduction in the amount of available CO_2 and counter the hindrance to its ability to photosynthesize and grow. *LCIB*, a gene upregulated in such conditions, encodes a protein potentially involved in uptake of CO_2 into the cell and in preventing the leakage of CO_2 out of the cell. This protein is indispensable for growth in air-level CO_2 (~350-400 ppm), since mutants in this gene are unable to grow (hence they are called *air dier* mutants). Several mutants that have second-site alterations that restore growth in air-level CO_2 (i.e., *suppress* the *air dier* phenotype) have been isolated. Identifying the genes that are mutated in these *suppressors* and the functions of the encoded proteins will help us better discern the role of *LCIB* and comprehend the workings of the entire mechanism.

To identify the locus of the mutated gene in one suppressor mutant, crosses were performed between the mutant strain and a non-mutant, polymorphic wild-type strain, and a large population of recombinant progeny that segregated *against* the mutation of interest was amassed. Using a strain that has unique single nucleotide polymorphisms (SNPs) as the non-mutant parent allowed us to seek a particular characteristic (the polymorphisms) in the region of interest in the genome of the progeny. With a sequenced genome, a library of SNPs in the polymorphic strain, and a pool of the genomic DNA from the entire population, we mapped the mutation to a specific region of the genome and narrowed potential candidates down to a small number of genes. By cosegregation analysis, we were able to confirm one of the candidates, *LCI15*, as the implicated gene.

Preliminary functional analyses with semi-quantitative RT-PCR and Western immunoblots reveal the LCI15 protein as possibly playing an overarching role in regulation in the CCM and offer terms for discussing potential methods by which the lack of LCI15 might potentially mask the deleterious effects of the absence of LCIB.

CHAPTER 1. GENERAL INTRODUCTION AND LITERATURE REVIEW

Chlamydomonas reinhardtii

The Chlamydomonas genus comprises unicellular green algae and includes at least 500 different species, of which *Chlamydomonas reinhardtii* is one of the most commonly used in the laboratory as a research tool for studying fundamental cellular processes. Algae can be found on damp soil, in seawater and fresh water, on ocean beds, and even in the snow.

The average *C. reinhardtii* cell has an ovate or ellipsoid shape, and the cell size varies with growth stage and strain (Figure 1). Estimates of diameter and volume of wild-type cells are about 10 μm and 75–150 μm^3, respectively (Chioccioli *et al.*, 2014; Umen and Goodenough, 2001b). The cytoplasm of the cell is largely occupied by a cup-shaped chloroplast, with the nucleus housed within the "cupped" area of the cytoplasm. Within the chloroplast, there is at least one pyrenoid, a membrane-less microcompartment that houses 40% to 90% of Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase, the enzyme that catalyzes the first major step of the Calvin cycle in photosynthesis) in the cell, depending on cellular and environmental conditions (Borkhsenious *et al.*, 1998; Mitchell *et al.*, 2014).

Like other algae in this genus, *C. reinhardtii* has two flagella that facilitate cell motility, and flagellar proteins have certain sensory functions. These threadlike appendages extend from the anterior side of the cell, with the axoneme comprising microtubules arranged in a 9+2 outer+central doublet configuration (Harris, 2009a). An organelle known as the eyespot detects light direction and initiates phototactic behavioral responses. It comprises rows of carotenoid-filled granules in the chloroplast (with each

row buttressed by a thylakoid membrane) and channelrhodopsins in the plasma membrane (Engel *et al.*, 2015; Thompson *et al.*, 2017; Ueki *et al.*, 2016).

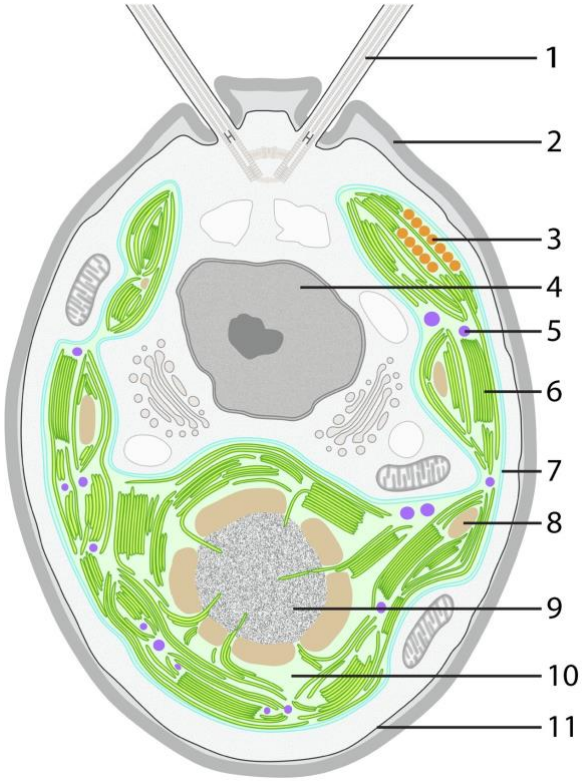


Figure 1. Ultrastructure of a *Chlamydomonas reinhardtii* cell (reprinted from Engel *et al.*, 2015). The cup-shaped chloroplast, shown here in color, occupies a large portion of the cytoplasm, and the nucleus (labeled "4") is housed within the "cupped" area. Other structures labeled here that are within (or part of) the chloroplast are the eyespot (3), a plastoglobule (5), thylakoids (6), the chloroplast envelope (7), a starch granule (8), the pyrenoid (9), and the chloroplast stroma (10). Cylindrical pyrenoid tubules continuous with granal thylakoids extend into the pyrenoid, penetrating the surrounding starch sheath. Also labeled are one of the two flagella (1), the cell wall (2), and the plasma membrane (11).

As was first proposed by Foster and Smyth (1980), and was later borne out experimentally (Kreimer and Melkonian, 1990; Matsunaga *et al.*, 2003; Ueki *et al.*,

2016), the carotenoids reflect and amplify the light striking it from the cell's outside onto the photoreceptor channelrhodopsin proteins, which can also capture photons when they initially strike the cell surface directly exterior to the eyespot. The carotenoid layers also absorb light striking them from their rear side (or the interior of the cell), thus shielding the photoreceptors from light coming through the cell body. This organelle also contains proteins that are purported to play a role in a number of cellular functions, including synthesis, transport, and degradation of other proteins in the cell (Schmidt *et al.*, 2006). Understanding flagellar motility, the proteins regulating it, and the co-ordination between the eyespot and flagella to achieve cellular phototaxis are areas of keen interest to researchers.

Cells also have a cell wall – a multilayered framework of glycoproteins – that offers structural fortification for the cell and controls movement of large molecules to or from the interior of the cell. Notably, in contrast to those in higher plants, *C. reinhardtii* cell walls do not contain cellulose. A number of cell wall-deficient mutants that have been isolated are used in transformation and mating experiments (Harris, 2009b, 2009d; Imam *et al.*, 1985). Apart from the ones discussed here, other plant cell organelles, including mitochondria, vacuoles, Golgi bodies, and endoplasmic reticuli, are seen in a typical cell.

C. reinhardtii is a photosynthetic autotroph but is also a facultative heterotroph, capable of growth in the dark when supplied with an alternate, organic carbon source such as acetate. This is particularly useful because mutants in the photosynthetic process can survive on the external carbon source and thus be studied. Cultures can be grown fairly easily in liquid or on solid (agar) media, with doubling time averaging about 7-8

hours. Reproduction occurs both asexually and sexually. During the haploid vegetative phase, cells reproduce asexually by mitosis, producing haploid daughter cells. When starved of nitrogen, vegetative cells undergo gametogenesis, differentiating into mating type *plus* or *minus* cells (*C. reinhardtii* is heterothallic, so the mating type of a strain is either *plus* or *minus*, i.e. mating type does not get arbitrarily assigned in each cell). The genotype is dictated by a single gene and follows Mendelian inheritance patterns. During mating, two isogametes of opposite mating types fuse to form a diploid zygote, which matures to a zygospore with a thick, protective wall. This remains dormant until the environment is favorable for germination following which it undergoes meiosis. The four haploid daughter cells (zoospores), two of each mating type, enter the vegetative phase. The tetrad of zoospores can be separated by careful dissection, a technique that comes in handy for genetic analyses of crosses. The tough fibrous wall of the zygospore renders it extremely robust, allowing it to endure harsh, unfavorable conditions. About 1-5% of zygotes undergo mitosis instead of germination and meiosis, resulting in vegetative diploids (Harris, 2009b, 2009d) that can be useful for studying recessive lethal mutations.

The *C. reinhardtii* nuclear genome, approximately 120 Mbp in 17 chromosomes, has been sequenced (Grossman *et al.*, 2003; Merchant *et al.*, 2007) and is published online (the U.S. Department of Energy–JGI *Chlamydomonas* sequencing project at www.phytozome.net/chlamy). The genome is largely GC-rich (64%) and comprises about 19,526 protein-coding transcripts. The available predicted transcriptome and proteome sequences are invaluable especially in efforts to understand possible gene function through forward and reverse genetic techniques. Sequences of the 203 kbp chloroplast and the 15.8 kbp mitochondrial genomes have also been determined, the latter

being the first in a photosynthetic organism to be fully sequenced (Buchanan *et al.*, 2015; Gray and Boer, 1988; Maul *et al.*, 2002; Michaelis *et al.*, 1990; Vahrenholz *et al.*, 1993). While transformation of all its three genomes has been performed successfully, *C. reinhardtii* is particularly amenable to nuclear and chloroplast transformation, making it an ideal organism for studying not just nuclear genes and their encoded proteins but also chloroplast biogenesis and function. In meiotic progeny, chloroplast genes are transmitted maternally and mitochondrial genes paternally, making this species an ideal model for studying how uniparental organelle inheritance is achieved and how related processes, such as selective destruction or protection of chloroplast/mitochondrial DNA, are regulated (Harris, 2009c; Umen and Goodenough, 2001a).

Photosynthesis and Rubisco

Photoautotrophic organisms have the unique ability to carry out photosynthesis – a process in which light energy; inorganic carbon compounds such as atmospheric CO_2; and water are utilized to produce carbohydrates, which can be used as structural components of the organism and as raw material for the production of energy in the form of ATP. Carbohydrates produced by photosynthesis are also a crucial sink for CO_2 in the Earth's atmosphere. These organisms are considered *producers* in the ecological food chain, since, through photosynthesis, they generate material that can become food used by *consumers* higher up in the chain.

Photosynthesis comprises a series of complex reactions which take place in largely two phases that can occur concurrently: light-dependent reactions and carbon cycle reactions. The light-dependent reactions, which take place in thylakoid membranes in chloroplasts, involve the conversion of light energy into chemical bond energy in the

form of ATP and NADPH, and the oxidation of water and production of O_2 as a byproduct. The carbon reactions (also known as the carbon reduction cycle, the Calvin-Benson-Bassham (CBC) cycle, or simply the Calvin cycle) occur in the chloroplast stroma, are light-independent, and synthesize carbohydrates from CO_2, using the ATP and NADPH produced in the light reactions as sources of energy.

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco), considered the most abundant protein on Earth (Ellis, 1979; Raven, 2013), is arguably also one of the most important, as it plays a major role in photosynthesis, and thus, in the sustenance of life on the planet. This enzyme is responsible for fixing CO_2 in the first step of the Calvin cycle, and has a fairly slow catalytic turnover rate of ~2–5 CO_2 molecules fixed per second (Bracher *et al.*, 2017). The fact that Rubisco exists in such copious amounts is explained not only by how vital it is to the existence of life but also perhaps by its inefficiency in catalyzing the carboxylation of ribulose-1,5-bisphosphate (RuBP) – the aforementioned CO_2-fixing, first step of the Calvin cycle – and by its participation in a competing oxygenation reaction as part of the photorespiratory pathway. Photorespiration is ostensibly a wasteful process as it uses energy that could be used in photosynthesis instead, and it also results in the loss of fixed carbon (although some of the carbon that enters the process does get recycled to be used in the Calvin cycle). Having first evolved during a time period when the Earth's atmosphere was essentially free of O_2 and when the ratio of CO_2 and O_2 was far greater than modern-day levels (current proportions are roughly 21% O_2 and 0.04% CO_2), Rubisco did not at first recognize O_2 as a substrate and catalyzed only CO_2. However, with the increase in atmospheric O_2 levels and the simultaneous decrease in CO_2 levels over time – due in large part to the advent of

oxygenic photosynthesis (a process that began in cyanobacteria) and the concomitant consumption of CO_2 and release of O_2 – the enzyme did not evolve to adequately distinguish between the two gases (Whitney *et al.*, 2011). Thus, in present day, it displays an affinity for the more-abundant O_2, favoring the oxidation of RuBP and the photorespiratory pathway.

Subsequent to the carboxylation of the 5-carbon RuBP molecule, the Calvin cycle yields two 3-carbon molecules of 3-phosphoglycerate (3-PGA). The 3-PGA is converted to glyceraldehyde-3-phosphate (G3P or GAP) after successive phosphorylation and reduction reactions, utilizing the energy in one ATP and one reductant NADPH. For every six G3P molecules generated, only one is used in the synthesis of sugars, amino acids, and fatty acids, while the remaining five are regenerated into RuBP. The result of the Calvin cycle is a net gain of six fixed carbon atoms with the usage of nine ATP and six NADPH molecules.

On the other hand, oxidation of RuBP immediately results in the formation of one molecule each of the 3-carbon 3-PGA and the 2-carbon 2-phosphoglycolate (2-PG). The latter inhibits carboxylase activity when present in large quantities (Falkowski and Raven, 2013) and is toxic to chloroplasts (Zelitch *et al.*, 2009). The 3-PGA molecule gets recycled to be used in the Calvin cycle, while, after a series of reactions, two 2-PG molecules are converted to one 3-PGA molecule (which can then enter the Calvin cycle) and one CO_2 molecule. During this process, CO_2 is released, and there is a net loss of three fixed carbon atoms. While in plants the photorespiratory pathway spans the chloroplast stroma, the peroxisome, and the mitochondria, in *C. reinhardtii,* it occurs

only in the stroma and mitochondria, with the peroxisomal reactions relocated to the latter organelle.

The CO_2 Concentrating Mechanism

To mitigate the effects of Rubisco's inefficient carboxylation activity and its incapacity to discriminate between CO_2 and O_2, some land plants have developed the well-studied C4 carbon metabolism and the Crassulacean acid metabolism (CAM). In aquatic photosynthetic organisms, another complicating factor to consider is that CO_2 diffuses about 10,000 times slower in water than it does in air (and HCO_3^- ions diffuse even more slowly), thus decelerating overall movement and fixation of CO_2 (Evans and Von Caemmerer, 1996; Ricklefs and Miller, 2000). Examples of these organisms are cyanobacteria and the green alga *Chlamydomonas reinhardtii* (hereafter "Chlamydomonas"), which have to obtain their inorganic carbon (Ci) (in either the form of CO_2, HCO_3^-, or the less common CO_3^{2-}) dissolved in their surroundings. They have developed CO_2-concentrating mechanisms (CCMs) which introduce a collection of changes in their gene expression, cellular structures, and physiology.

CCM Acclimation States

Vance and Spalding (2005) established that there are three separate ambient CO_2 levels based on how Chlamydomonas acclimates to them: high CO_2 (HCO_2; 1–5% CO_2), low CO_2 (LCO_2; 0.03–0.4% CO_2 or 300–4000 ppm CO_2), and very low CO_2 ($VLCO_2$; <0.02% CO_2 or <200 ppm CO_2). Cells acclimated to LCO_2 and $VLCO_2$ conditions (together referred to as 'limiting CO_2' conditions) employ certain measures for active Ci uptake, which leads to increased Ci accumulation relative to that seen in cells grown in

HCO_2 conditions. An increase in Ci accumulation subsequently gives rise to increased photosynthetic affinity and decreased $K_{1/2}$ for CO_2 (Duanmu *et al.*, 2009a; Moroney and Tolbert, 1985; Spalding *et al.*, 1983a, 1983b, 1983c).

At the molecular level, different elements that compensate for the limiting CO_2 conditions have been found to exist in LCO_2 and $VLCO_2$ (Wang and Spalding, 2014a). Compared to LCO_2 cells, those from $VLCO_2$ cultures have a higher affinity for CO_2 (thus a lower $K_{1/2}$) and a lower V_{max}. They also have a longer cell-doubling time, smaller average cell size, and lower chlorophyll content than what is observed in LCO_2-grown cells (Vance and Spalding, 2005). Various studies of gene expression at the transcript level (Brueggeman *et al.*, 2012; Fang *et al.*, 2012; Im *et al.*, 2003; Miura *et al.*, 2004; Yamano and Fukuzawa, 2009; Yamano *et al.*, 2008) have revealed upregulation or induction of genes associated with the CCM, and have been invaluable in genetic, molecular, and physiological experiments for understanding protein function in various acclimation states. The relationship between CCM proteins and the form of Ci they might have an affinity towards can also be studied by varying the ambient pH, which partially dictates the ratio of CO_2 and HCO_3^- available to the cells.

There are three main essential components in the algal CCM: the pyrenoid, an area inside the chloroplast where Rubisco is highly concentrated (so that large amounts of CO_2 can be delivered here and be fixed by the enzyme); transporters that ferry Ci into the cell and increase its overall intracellular concentration; and carbonic anhydrases that facilitate conversion between different forms of inorganic carbon (Ci; CO_2, HCO_3^-, CO_3^{2-}) in various locations in the cell.

The Pyrenoid

The sequestration of Rubisco within a cellular region where CO_2 can be made readily available to the enzyme forces the carboxylase activity of Rubisco to take precedence over the oxygenase activity and improves the *rate* at which carboxylation occurs (Mann, 1999). In prokaryotes such as cyanobacteria, these regions – called carboxysomes – contain Rubisco and a carbonic anhydrase that facilitates rapid conversion of HCO_3^- to CO_2. The physical structure of carboxysomes is defined by semipermeable protein shells that, along with the pH levels inside and outside, allow for the inward diffusion of HCO_3^- but limit the outward movement of the CO_2 that is formed within (Badger and Price, 2003; Bracher *et al.*, 2017; Espie and Kimber, 2011; Giordano *et al.*, 2005; Kerfeld *et al.*, 2010).

In eukaryotic algae that employ CCMs, these regions are called pyrenoids and are located within the chloroplast stroma. They are not bound by a membrane and comprise a matrix of tubules that are continuous with thylakoid tubules, as seen in Figure 1 (Engel *et al.*, 2015; Markelova *et al.*, 2009; Rochaix, 2017). In high CO_2 environments, only about half the Rubisco that exists in the cell is located in the pyrenoid, but in limiting CO_2 conditions, as much as 90% of the enzyme is found in it (Borkhsenious *et al.*, 1998; Mitchell *et al.*, 2014).

While pyrenoids themselves are essential for successful establishment of the CCM (Fukuzawa *et al.*, 2001; Genkov *et al.*, 2010; Ma *et al.*, 2011; Mitchell *et al.*, 2017), proper formation of these microcompartments has been found to be dependent on some other proteins. CIA6 is one such protein, as is observed in *CIA6* mutants which form smaller pyrenoids to which Rubisco does not localize and which exhibit impaired growth in LCO_2 (Ma *et al.*, 2011). Both the large and small subunits of Rubisco are also critical

to pyrenoid establishment, and the small subunits appear to be important for directing Rubisco to the pyrenoid (Genkov *et al.*, 2010; Meyer *et al.*, 2012; Rawat *et al.*, 1996). Mackinder *et al.* (2016) recently discovered EPYC1 – a protein present in the pyrenoid in amounts almost equal to those of Rubisco – binds the enzyme to form the pyrenoid matrix and is important in various aspects of pyrenoid formation and structure.

Ci Transporters

To accumulate Ci in the cell and ultimately in the pyrenoid, Ci transporter systems that facilitate active transport across the plasma membrane and the chloroplast envelope are essential.

Putative Ci transporters in the plasma membrane

In the plasma membrane, two proteins have been recognized as being candidate transporters: HLA3 (Im and Grossman, 2002; Miura *et al.*, 2004; Yamano *et al.*, 2015) and LCI1 (Burow *et al.*, 1996; Ohnishi *et al.*, 2010). HLA3, an ABC (ATP-binding cassette)-type transporter, is regulated by CIA5 (a protein identified as the master regulator of CCM gene expression in low CO_2 conditions (Fukuzawa *et al.*, 2001; Miura *et al.*, 2004; Xiang *et al.*, 2001)) and is expressed in high light and low CO_2 (but not under high light and high CO_2). Duanmu *et al.* (2009a) showed that when HLA3 is knocked down in wild-type cells using RNAi, there is a decrease in Ci uptake at high pH, which was later confirmed by Yamano *et al.* (2015) with HLA3 insertional mutants. Since HCO_3^- is the favored form of Ci at alkaline pH, it is likely that HLA3 is involved in transporting HCO_3^-.

LCI1 is also proposed to be controlled by CIA5 and LCR1, a Myb-DNA binding transcription factor (Yoshioka *et al.*, 2004). In LCO_2 conditions, an increase in photosynthetic Ci affinity and uptake was observed in an *LCR1* mutant overexpressing *LCI1* when compared to an *LCR1* mutant that has very low expression of *LCI1* (Ohnishi *et al.*, 2010). Transgenic *LCI1* expression in HCO_2 conditions led to increased Ci accumulation. It appears that LCI1 might be associated with Ci uptake, however, whether LCI1 transports/accumulates CO_2 or HCO_3^- has not been clearly elucidated.

Interestingly, Mackinder *et al*. (2017) discovered that HLA3 and LCI1 interact with each other and form a complex with a P-type ATPase in the plasma membrane. They theorize that the ATPase might be facilitating the maintenance of an H^+ gradient that in turn helps HLA3 and/or LCI1 mediate HCO_3^- uptake. The association of ATPase and LCI1 with the plasma membrane is consistent with previous results (Norling *et al.*, 1996; Ohnishi *et al.*, 2010; Yamano *et al.*, 2015).

Putative Ci transporters in the chloroplast envelope

Another barrier for the entry of Ci into the cell is the chloroplast envelope, which houses at least one confirmed protein, LCIA (CCP1 and CCP2, earlier believed to be chloroplast envelope proteins (Mason *et al.*, 1990; Ramazanov *et al.*, 1993), were recently localized to the mitochondria by Atkinson *et al*. (2016)). LCIA, part of the nitrate assimilation-related (NAR) protein family (but regulated by CO_2 and not nitrogen (Miura *et al.*, 2004)), is reported to be a HCO_3^- transporter. While RNAi-knockdown of *HLA3* alone resulted in some decrease in Ci affinity, co-knockdown of *HLA3* and *LCIA* led to more substantial negative effects on Ci accumulation, photosynthetic Ci affinity, and overall growth rate in high pH, HCO_3^--rich conditions in $VLCO_2$ (Duanmu *et al.*,

2009a). Gao *et al*. (2015) observed that overexpression of *HLA3* and *LCIA* resulted in an increase in photosynthetic O_2 evolution rate in the $VLCO_2$ range of Ci but not at higher Ci levels, supporting earlier observations. *HLA3*, *LCIA*, and *HLA3-LCIA* single and double mutants, and lines overexpressing these genes, also showed that HLA3 and LCIA work in tandem in HCO_3^- uptake in the cell (Yamano *et al.*, 2015).

Ci accumulation in the chloroplast stroma

LCIB, a plastid protein, forms a hetero-hexameric complex with LCIC (Wang and Spalding, 2014b; Yamano *et al.*, 2010), a protein from the LCIB family. This family comprises four homologous proteins: LCIB, LCIC, LCID, and LCIE. *LCIB* and *LCIC* have been shown to have the highest expression levels in limiting CO_2 conditions, while *LCIE* has far lower expression than *LCID*. *LCIC* mutants do not reveal a distinct growth phenotype, and accumulate a homohexameric LCIB complex, suggesting that LCIC is not essential for LCIB function or complex formation (Wang and Spalding, unpublished observations). However, it was observed that in $VLCO_2$-acclimated *LCIC* mutant cells, the LCIB/LCIC complex failed to migrate to the peri-pyrenoid region, in contrast with what is observed in wild-type cells, suggesting that LCIC might play a role in localization of the complex rather than directly in its function.

LCIB is postulated to play an important role in the uptake of CO_2 and in preventing its leakage from the chloroplast in limiting CO_2 conditions. It has been localized to the periphery of the pyrenoid in $VLCO_2$ but stays dispersed throughout the chloroplast in HCO_2 or LCO_2 (Duanmu *et al.*, 2009b; Wang and Spalding, 2014a, 2014b). *LCIB* mutants (Spalding *et al.*, 1983c; Wang and Spalding, 2006) experience significant loss of accumulated CO_2 that is unused by Rubisco (which is located in the

pyrenoid) because CO_2 can easily diffuse out of the cell. While these mutant cell lines are unable to survive in "air level" CO_2 (about 300–450 ppm, a sub-range of LCO_2) – and are hence called "*air diers*" – they exhibit near-wild type level growth and photosynthesis in $VLCO_2$. This provided evidence that LCIB played a critical role in LCO_2. Wang and Spalding (2014a) found that LCIB plays a role not only in LCO_2 but in $VLCO_2$ as well.

In LCO_2, it is surmised that LCIB stops the outflow of CO_2 – either the CO_2 that is generated in the pyrenoid via dehydration of HCO_3^- by the pyrenoid-localized carbonic anhydrase CAH3, or that which enters the cell from the outside – from the cell, thus contributing to the accumulation of stromal Ci in the form of HCO_3^-, which can then be fed (or fed back) to CAH3 for dehydration, and subsequent delivery to Rubisco in the pyrenoid (Duanmu *et al.*, 2009b; Wang and Spalding, 2014b). It has been proposed that LCIB might function as a unidirectional CA, catalyzing the hydration of CO_2 to HCO_3^- (Duanmu *et al.*, 2009b) in the chloroplast stroma, and this was supported by recent work by Jin *et al*. (2016) which revealed that some structural features of LCIB were similar to those of β-CAs. HCO_3^- uptake involving HLA3 and LCIA is inhibited in this CO_2 level (Wang and Spalding, 2014a; Yamano *et al.*, 2015).

In $VLCO_2$, LCIB works in concert with the HCO_3^- uptake system (involving HLA3 and LCIA), which is more active than the LCIB-associated CO_2 uptake system for Ci accumulation in this CO_2 level. LCIB still likely captures CO_2 already inside the cell and prevents its leakage to the outside, similar to its function in LCO_2. The complementary functions of these two pathways in $VLCO_2$ is supported by RNAi-knockdown of *HLA3* in *LCIB* mutants resulting in hampered growth (Duanmu *et al.*, 2009a).

Other proteins

CCP1 and CCP2 (chloroplast carrier proteins 1 and 2) have shown 95.7% identity in their polypeptide sequences and have six predicted transmembrane domains, which, along with some chloroplast and envelope membrane preparations, earlier implied that these two proteins are located in the chloroplast envelope (Chen *et al.*, 1997; Geraghty *et al.*, 1990; Mason *et al.*, 1990; Ramazanov *et al.*, 1993; Spalding and Jeffrey, 1989). However, they were recently located in the mitochondria by Atkinson *et al*. (2016). They are also among the most highly expressed proteins under limiting CO_2 conditions, and although RNAi-mediated knockdown of *CCP1/2* resulted in defective growth in $VLCO_2$, it did not have much of an effect on photosynthetic activity (Pollock *et al.*, 2004). These data suggest that mitochondria and mitochondrial membrane transport are important for growth-related, rather than photosynthetic, activity in the CCM (Atkinson *et al.*, 2016). It has been speculated that there is another system or set of genes that masks any deleterious effects of *CCP1/2* mutations, as the lack of a considerable effect in $VLCO_2$ suggests (Pollock *et al.*, 2004).

Carbonic Anhydrases

CO_2 is the only Ci form that Rubisco can use as a substrate, so when the cells obtain HCO_3^- from the surroundings, it has to be dehydrated before it can serve as substrate for Rubisco. Carbonic anhydrases (CAs or CAHs) are enzymes that facilitate the conversion of HCO_3^- and H^+ ions to CO_2 and water (and vice-versa).

CAH1, a CA in the periplasm, may be involved in maintaining a supply of HCO_3^- or CO_2 for the respective transporters to ferry into the cell. It is induced in limiting CO_2 conditions in the presence of light, while CAH2, likely a CAH1 isozyme with 91.8%

identity between their amino acid sequences, is induced when exposed to HCO_2 conditions and light (Coleman and Grossman, 1984; Fujiwara *et al.*, 1990; Fukuzawa *et al.*, 1990a; Fukuzawa *et al.*, 1990b). *CAH1* mutants do not have any significant growth phenotype and show only a small decrease in photosynthetic Ci affinity in $VLCO_2$ (Van and Spalding, 1999; Van *et al.*, 2001). It is possible that there are compensating elements that mask the effects of a CAH1 absence, especially considering that more than one Ci uptake pathway exists in $VLCO_2$ conditions (Wang and Spalding, 2014a).

CAH3, a CA localized to the thylakoid lumen (Karlsson *et al.*, 1998), is likely dehydrating incoming HCO_3^- to CO_2 that can be provided as substrate for carboxylation by Rubisco in the pyrenoid. Loss-of-function *CAH3* mutations result in deteriorating growth with decreasing levels of ambient CO_2 (poor growth in LCO_2 and no growth in $VLCO_2$), while there is a large internal accumulation of HCO_3^- (Spalding *et al.*, 1983a). *LCIB-CAH3* double mutants also show a similar phenotype, suggesting that CAH3 probably functions upstream of LCIB. A *CAH3* mutation can restore growth in LCO_2 in an *LCIB (air dier)* mutant (Duanmu *et al.*, 2009b). The inability of *LCIB* mutants to accumulate internal Ci, and the over-accumulation of Ci in cells in LCO_2 in *CAH3* mutants as well as *LCIB-CAH3* mutants is further evidence of LCIB itself not being directly involved in Ci transport but rather having a role in the recapture of the CO_2 that is made by CAH3 dehydrating HCO_3^- and in the capture of CO_2 that enters the stroma from outside the cell. It has been reported that CAH3 is associated with the PSII (photosystem II) machinery in the thylakoids (Benlloch *et al.*, 2015; Villarejo *et al.*, 2002). During limiting CO_2 acclimation, most of the CAH3 is activated by phosphorylation and relocates to the thylakoids associated with the pyrenoid (Blanco-Rivero *et al.*, 2012).

Considering the localization of the LCIB/LCIC complex around the pyrenoid in VLCO$_2$ environments, it has been suggested that the complex might function as a physical barricade that prevents leakage of CAH3-generated CO_2 away from the pyrenoid (Duanmu *et al.*, 2009b; Yamano *et al.*, 2010).

CAH4 and CAH5, localized to the mitochondria (Eriksson *et al.*, 1995; Eriksson *et al.*, 1996; Mackinder *et al.*, 2017), differ by only a single amino acid (Moroney *et al.*, 2011) and are highly expressed in limiting CO_2 conditions. Given that, in HCO$_2$ cells, the mitochondria are in the cup of the chloroplast and those in limiting CO_2 cells migrate close to the plasma membrane (Geraghty and Spalding, 1996), it is possible CAH4/5 are involved in quickly providing energy to transport proteins in the plasma membrane, or in offsetting the effects of the surge in photorespiratory activity that is observed during acclimation to limiting CO_2 (Spalding, 2009).

The high pH in the chloroplast stroma favors accumulation of HCO_3^-, and CAH6, which was reported to be localized in the stroma (Mitra *et al.*, 2004) was earlier suggested to carry out the conversion of the acquired stromal CO_2 to HCO_3^-. However, Mackinder *et al.* (2017) recently discovered that CAH6 localizes abundantly to the flagella, surmising that there might still be some CAH6 in the stroma, but in very small amounts. Based on data from previous work (Choi *et al.*, 2016; Hu *et al.*, 2010), they also suggested that CAH6 might be an inorganic carbon sensor in the flagella, facilitating chemotaxis of the cell towards HCO_3^-.

CAH7 and CAH8 share 63% similarity and are constitutively expressed (Moroney and Ynalvez, 2007). The location of CAH7 was not demonstrated by immunolocalization experiments, but it is predicted to likely be in the chloroplast, and CAH8 was found to be

in the periplasmic space (Ynalvez *et al.*, 2008). This is particularly interesting because of the lack of a distinct phenotype in the *CAH1* mutant; it is possible that CAH8 compensates, at least partially, for the absence of CAH1.

CAH9, predicted to be in the cytoplasm, was found to be closely related to bacterial CAs (Moroney and Ynalvez, 2007). It has not been well-characterized as yet and its function is still unknown.

Genes that encode three more putative CAs have been identified in Chlamydomonas. Unlike the others CAs, which are either alpha- or beta-CAs, these are gamma-CAs, which are associated with Complex I of the mitochondrial electron transport chain in algae and higher plants (Klodmann *et al.*, 2010; Sunderhaus *et al.*, 2006). They appear to be fairly well-expressed at the transcript level, but their abundance at the protein level is unknown (Moroney *et al.*, 2011).

Regulation of the CCM

Transcriptome analyses by Fang *et al.* (2012) and Brueggeman *et al.* (2012) have revealed clear changes in gene expression in LCO_2 or $VLCO_2$ levels of growth, either by induction of genes not normally expressed in HCO_2 conditions or by upregulation of genes that are expressed constitutively but at a lower level in HCO_2. Interestingly, there were no genes detected that showed differential expression between the limiting CO_2 levels, LCO_2 and $VLCO_2$, even though there are clear differences between the physiological responses of cells in these two acclimation states.

CIA5 (also known as CCM1) is considered a “master regulator” of CCM genes, since in *CIA5* mutants, there is a clear decrease in transcript abundance of genes that are normally induced or upregulated in wild type cells upon exposure to limiting CO_2

conditions (Fang *et al.*, 2012; Miura *et al.*, 2004; Moroney *et al.*, 1989). *CIA5* mutants exhibit slow growth in limiting CO_2: while their growth is comparable to that of wild-type cells in HCO_2, it deteriorates in LCO_2, and cannot be sustained in $VLCO_2$ (Moroney *et al.*, 1989). In addition, these mutants are unable to form a pyrenoid and have low Ci affinity and accumulation in limiting CO_2 (Fukuzawa *et al.*, 1998; Fukuzawa *et al.*, 2001). *CIA5* itself is constitutively expressed at a low abundance level through all three CO_2 conditions while its regulatory effects are seen only in limiting CO_2, so it is possible that CIA5 is activated at a post-translational stage in order for it to perform as a transcriptional regulator of CCM genes (Fang *et al.*, 2012; Miura *et al.*, 2004; Wang *et al.*, 2005; Xiang *et al.*, 2001). CIA5 has putative zinc-finger and DNA-binding domains (Kohinata *et al.*, 2008; Somanchi and Moroney, 1999), and Chen *et al.* (2017) recently discovered that a mini-CIA5 construct with a partial activation domain and the putative zinc-binding motif was sufficient to complement the *CIA5* mutation, restoring growth and CCM gene expression levels to those seen in *CIA5* mutants complemented with the full CIA5 construct or in wild type cells. LCR1, a Myb transcription factor, was found to regulate *CAH1, LCI1,* and *LCI6* expression in limiting CO_2, and is itself induced in limiting CO_2 conditions and is regulated by CIA5 (Yoshioka *et al.*, 2004).

CAS, a calcium (Ca^{2+})-binding protein, was recently discovered as being a regulator of the CCM (Wang *et al.*, 2016). *CAS* mutants exhibited decreased transcript levels of a number of LCO_2-induced genes, including *HLA3* and *LCIA*. The CAS-mediated regulation of these HCO_3^- transporters might be occurring through signaling by the influx of Ca^{2+} into the cells. This protein re-localizes from the thylakoid membrane in non-CCM conditions to the pyrenoid when the CCM is induced, and the loss of CAS

might result in the structural impairment of the thylakoids or the pyrenoid in LCO$_2$. CIA5 and CAS appear to regulate the CCM independently and in parallel, since a decrease in the level of one protein did not affect that of the other.

Broad Overview of the CCM

The proposed model of the CCM that is induced in Chlamydomonas is illustrated in Figure 2. With its ultimate goal being that of sequestering CO_2 near pyrenoidal Rubisco, the CCM is understood as beginning, functionally, with the periplasmic CA, CAH1, which possibly catalyzes interconversion of the two common Ci species, HCO_3^- and CO_2, with the CO_2 then diffusing through the plasma membrane and the chloroplast envelope, and the HCO_3^- being actively transported across these two barriers. The HLA3/LCI1 complex aids in this active transport of HCO_3^- across the plasma membrane, while LCIA mediates passage of HCO_3^- through the chloroplast envelope. Once in the chloroplast stroma, accumulated HCO_3^- is possibly being transported across the thylakoid membrane by a hypothetical transporter and delivered to the thylakoid lumen, while CO_2, possibly being hydrated to HCO_3^- by the putative unidirectional CA activity of the LCIB/LCIC complex, diffuses through. In the thylakoid lumen, HCO_3^- travels to CAH3 within the pyrenoid-associated thylakoid tubules and is dehydrated to CO_2, which subsequently diffuses from those tubules into the pyrenoid and is fixed by Rubisco. However, some CO_2 that escapes Rubisco and diffuses back into the chloroplast stroma is believed to be rehydrated to HCO_3^- (possibly by the aforementioned putative unidirectional CA activity of LCIB/LCIC or a different protein), which is then redirected from the stroma to the thylakoid lumen. This forestalls the easy, diffusive leakage of intracellular CO_2 to the exterior of the cell.

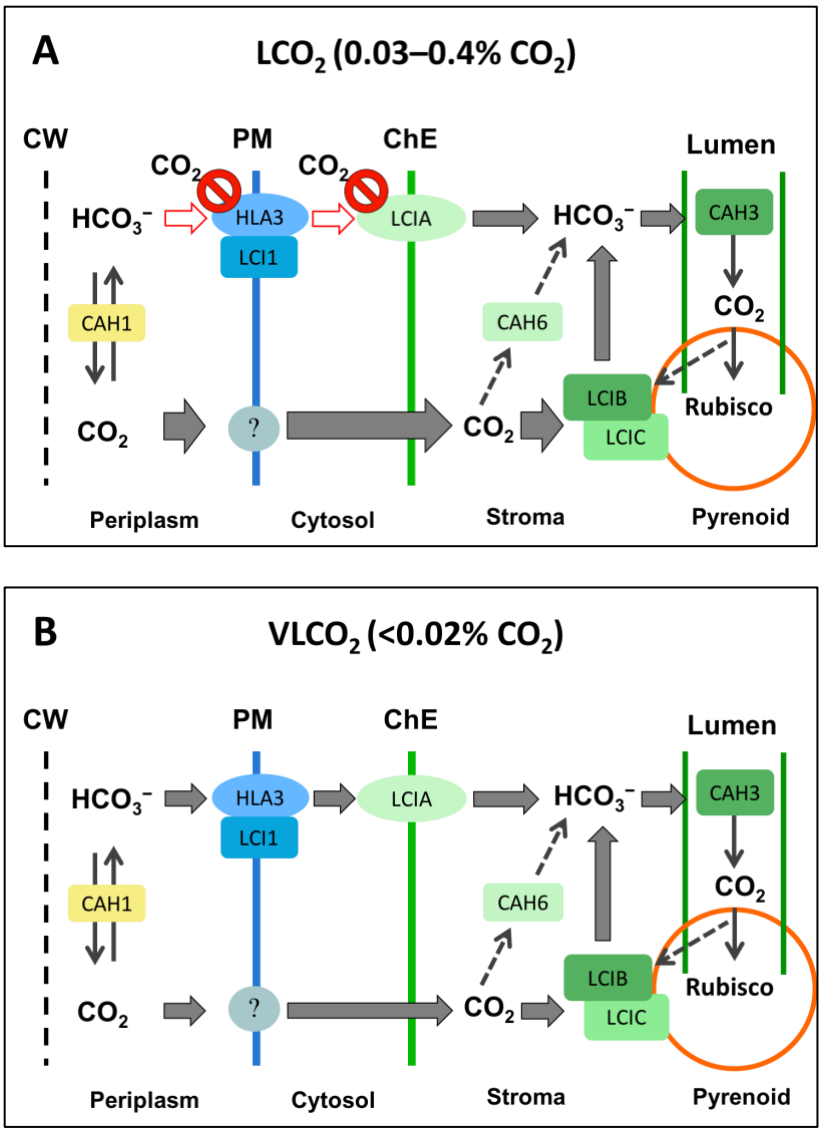


Figure 2. The hypothetical model of the CCM as illustrated in Wang and Spalding (2014a), with the mechanisms of Ci uptake proposed in (A) LCO_2 conditions (0.03–0.4% CO_2 or 300–4000 ppm) and (B) $VLCO_2$ conditions (<0.02% CO_2 or <200 ppm).

As proposed in Wang and Spalding (2014a), in $VLCO_2$ conditions (Figure 2B), HCO_3^- is transported to the chloroplast stroma by LCIA and related proteins, and CO_2 is converted to HCO_3^- by LCIB to add to the stromal HCO_3^- pool. Here, both LCIA and LCIB are involved in accumulation of internal Ci. In contrast, in LCO_2 conditions (Figure

2A), LCIB is largely responsible for Ci accumulation, while LCIA and related proteins are inhibited by CO_2. This explains the indispensable nature of LCIB in LCO_2-acclimated cells and the reason *LCIB* mutants cannot sustain growth in this CO_2 level.

Dissertation Organization

Second-site suppressors of *LCIB/air dier* mutants were generated earlier (Duanmu and Spalding, 2011), and one of these suppressor mutations, *su1*, is the focus of this dissertation. Chapter 2 covers the identification of the *SU1* suppressor locus; this was achieved by deep sequencing pooled DNA from a collection of progeny from crosses of the strain carrying the mutation of interest (*pmp-su1*) and a polymorphic strain. *LCI15* was then confirmed genetically as being the implicated gene. In Chapter 3, the possible mechanism by which LCI15 facilitates suppression is explored by studying the transcript and protein expression levels of several other CCM genes. Chapter 4 summarizes and concludes the dissertation.

References

Atkinson, N., Feike, D., Mackinder, L.C., 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.

Badger, M.R., & Price, G.D. 2003. CO_2 concentrating mechanisms in cyanobacteria: molecular components, their diversity and evolution. *J Exp Bot* 54:609-622.

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* 167:950-962.

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.

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* 116:1585-1591.

Bracher, A., Whitney, S.M., Hartl, F.U., & Hayer-Hartl, M. 2017. Biogenesis and Metabolic Maintenance of Rubisco. *Annu Rev Plant Biol* 68:29-60.

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.

Buchanan, B.B., Gruissem, W., & Jones, R.L. 2015. Biochemistry and Molecular Biology of Plants, 2nd ed. John Wiley and Sons.

Burow, M.D., Chen, Z.Y., Mouton, T.M., & Moroney, J.V. 1996. Isolation of cDNA clones of genes induced upon transfer of *Chlamydomonas reinhardtii* cells to low CO_2. *Plant Mol Biol* 31:443-448.

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 Research* 24, Part A:207-217.

Chen, Z.Y., Lavigne, L.L., Mason, C.B., & Moroney, J.V. 1997. Cloning and overexpression of two cDNAs encoding the low-CO_2-inducible chloroplast envelope protein LIP-36 from *Chlamydomonas reinhardtii*. *Plant Physiol* 114:265-273.

Chioccioli, M., Hankamer, B., & Ross, I.L. 2014. Flow cytometry pulse width data enables rapid and sensitive estimation of biomass dry weight in the microalgae *Chlamydomonas reinhardtii* and *Chlorella vulgaris*. *PLoS One* 9:e97269.

Choi, H.I., Kim, J.Y., Kwak, H.S., Sung, Y.J., & Sim, S.J. 2016. Quantitative analysis of the chemotaxis of a green alga, *Chlamydomonas reinhardtii*, to bicarbonate using diffusion-based microfluidic device. *Biomicrofluidics* 10:014121.

Coleman, J.R., & Grossman, A.R. 1984. Biosynthesis of carbonic anhydrase in *Chlamydomonas reinhardtii* during adaptation to low CO_2. *Proc Natl Acad Sci U S A* 81:6049-6053.

Duanmu, D., Miller, A.R., Horken, K.M., Weeks, D.P., & Spalding, M.H. 2009a. Knockdown of limiting-CO_2–induced gene *HLA3* decreases HCO_3^- transport and photosynthetic Ci affinity in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 106:5990-5995.

Duanmu, D., & Spalding, M.H. 2011. Insertional suppressors of *Chlamydomonas reinhardtii* that restore growth of air-dier *lcib* mutants in low CO_2. *Photosynth Res* 109:123-132.

Duanmu, D., Wang, Y., & Spalding, M.H. 2009b. Thylakoid lumen carbonic anhydrase (*CAH3*) mutation suppresses air-dier phenotype of *LCIB* mutant in *Chlamydomonas reinhardtii*. *Plant Physiol* 149:929-937.

Ellis, R.J. 1979. The most abundant protein in the world. *Trends in Biochemical Sciences* 4:241-244.

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:e04889.

Eriksson, M., Gardestrom, P., & Samuelsson, G. 1995. Isolation, Purification, and Characterization of Mitochondria from *Chlamydomonas reinhardtii*. *Plant Physiol* 107:479-483.

Eriksson, M., Karlsson, J., Ramazanov, Z., Gardeström, P., & Samuelsson, G. 1996. Discovery of an algal mitochondrial carbonic anhydrase: molecular cloning and characterization of a low-CO_2–induced polypeptide in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 93:12031-12034.

Espie, G.S., & Kimber, M.S. 2011. Carboxysomes: cyanobacterial RubisCO comes in small packages. *Photosynth Res* 109:7-20.

Evans, J.R., & Von Caemmerer, S. 1996. Carbon Dioxide Diffusion inside Leaves. *Plant Physiology* 110:339-346.

Falkowski, P.G., & Raven, J.A. 2013. Aquatic Photosynthesis: Second Edition. Princeton University Press.

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.

Foster, K.W., & Smyth, R.D. 1980. Light Antennas in phototactic algae. *Microbiol Rev* 44:572-630.

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.

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

Fukuzawa, H., Fujiwara, S., Yamamoto, Y., Dionisio-Sese, M.L., & Miyachi, S. 1990b. cDNA cloning, sequence, and expression of carbonic anhydrase in *Chlamydomonas reinhardtii*: regulation by environmental CO_2 concentration. *Proc Natl Acad Sci U S A* 87:4383-4387.

Fukuzawa, H., Ishizaki, K., Miura, K., Matsueda, S., Ino-ue, T., Kucho, K.-i., & Ohyama, K. 1998. Isolation and characterization of high-CO_2 requiring mutants from *Chlamydomonas reinhardtii* by gene tagging. *Canadian journal of botany* 76:1092-1097.

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.

Gao, H., Wang, Y., Fei, X., Wright, D.A., & Spalding, M.H. 2015. Expression activation and functional analysis of HLA3, a putative inorganic carbon transporter in *Chlamydomonas reinhardtii*. *Plant J* 82:1-11.

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* 285:19833-19841.

Geraghty, A.M., Anderson, J.C., & Spalding, M.H. 1990. A 36 Kilodalton Limiting-CO_2 Induced Polypeptide of *Chlamydomonas* Is Distinct from the 37 Kilodalton Periplasmic Carbonic Anhydrase. *Plant Physiol* 93:116-121.

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.

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.

Gray, M.W., & Boer, P.H. 1988. Organization and expression of algal (*Chlamydomonas reinhardtii*) mitochondrial DNA. *Philos Trans R Soc Lond B Biol Sci* 319:135-147.

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.

Harris, E.H. 2009a. Chapter 4 - Motility and Behavior A2. Pp. 89-117 in: Stern, D.B., & Witman, G.B., (eds), *The Chlamydomonas Sourcebook (Second Edition)*. Academic Press, London. p 89-117.

Harris, E.H. 2009b. Chapter 5 - The Sexual Cycle A2. Pp. 119-157 in: Stern, D.B., & Witman, G.B., (eds), *The Chlamydomonas Sourcebook (Second Edition)*. Academic Press, London. p 119-157.

Harris, E.H. 2009c. Chapter 7 - Organelle Heredity A2. Pp. 211-240 in: Stern, D.B., & Witman, G.B., (eds), *The Chlamydomonas Sourcebook (Second Edition)*. Academic Press, London. p 211-240.

Harris, E.H. 2009d. Chapter 8 - *Chlamydomonas* in the Laboratory A2. Pp. 241-302 in: Stern, D.B., & Witman, G.B., (eds), *The Chlamydomonas Sourcebook (Second Edition)*. Academic Press, London. p 241-302.

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; sup pp 81-18.

Im, C.S., & Grossman, A.R. 2002. Identification and regulation of high light-induced genes in *Chlamydomonas reinhardtii*. *Plant J* 30:301-313.

Im, C.S., Zhang, Z., Shrager, J., Chang, C.W., & Grossman, A.R. 2003. Analysis of light and CO_2 regulation in *Chlamydomonas reinhardtii* using genome-wide approaches. *Photosynth Res* 75:111-125.

Imam, S.H., Buchanan, M.J., Shin, H.C., & Snell, W.J. 1985. The *Chlamydomonas* cell wall: characterization of the wall framework. *J Cell Biol* 101:1599-1607.

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 U S A* 113:14716-14721.

Karlsson, J., Clarke, A.K., Chen, Z.Y., Hugghins, S.Y., Park, Y.I., Husic, H.D., Moroney, J.V., & Samuelsson, G. 1998. A 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.

Kerfeld, C.A., Heinhorst, S., & Cannon, G.C. 2010. Bacterial microcompartments. *Annu Rev Microbiol* 64:391-408.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jänsch, L., & Braun, H.P. 2010. Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22:797-810.

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.

Kreimer, G., & Melkonian, M. 1990. Reflection confocal laser scanning microscopy of eyespots in flagellated green algae. *Eur J Cell Biol* 53:101-111.

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.

Mackinder, L.C., 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 U S A* 113:5958-5963.

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.e114.

Mann, C.C. 1999. Genetic engineers aim to soup up crop photosynthesis. *Science* 283:314-316.

Markelova, A.G., Sinetova, M.P., Kupriyanova, E.V., & Pronina, N.A. 2009. Distribution and functional role of carbonic anhydrase Cah3 associated with thylakoid membranes in the chloroplast and pyrenoid of *Chlamydomonas reinhardtii*. *Russian Journal of Plant Physiology* 56:761.

Mason, C.B., Manuel, L.J., & Moroney, J.V. 1990. A New Chloroplast Protein Is Induced by Growth on Low CO_2 in *Chlamydomonas reinhardtii*. *Plant Physiol* 93:833-836.

Matsunaga, S., Watanabe, S., Sakaushi, S., Miyamura, S., & Hori, T. 2003. Screening effect diverts the swimming directions from diaphototactic to positive phototactic in a disk-shaped green flagellate *Mesostigma viride*. *Photochem Photobiol* 77:324-332.

Maul, J.E., Lilly, J.W., Cui, L., dePamphilis, C.W., Miller, W., Harris, E.H., & Stern, D.B. 2002. The *Chlamydomonas reinhardtii* plastid chromosome: islands of genes in a sea of repeats. *Plant Cell* 14:2659-2679.

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., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W.C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., & Grossman, A.R. 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* 318:245-250.

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 U S A* 109:19474-19479.

Michaelis, G., Vahrenholz, C., & Pratje, E. 1990. Mitochondrial DNA of *Chlamydomonas reinhardtii*: the gene for apocytochrome b and the complete functional map of the 15.8 kb DNA. *Mol Gen Genet* 223:211-216.

Mitchell, M.C., Metodieva, G., Metodiev, M.V., Griffiths, H., & Meyer, M.T. 2017. Pyrenoid loss impairs carbon-concentrating mechanism induction and alters primary metabolism in *Chlamydomonas reinhardtii*. *J Exp Bot* 68:3891-3902.

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* 166:1073-1082.

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.

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* 135:1595-1607.

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.

Moroney, J.V., Ma, Y., Frey, W.D., Fusilier, K.A., Pham, T.T., Simms, T.A., DiMario, R.J., Yang, J., & Mukherjee, B. 2011. The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. *Photosynth Res* 109:133-149.

Moroney, J.V., & Tolbert, N.E. 1985. Inorganic Carbon Uptake by *Chlamydomonas reinhardtii*. *Plant Physiol* 77:253-258.

Moroney, J.V., & Ynalvez, R.A. 2007. Proposed carbon dioxide concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryot Cell* 6:1251-1259.

Norling, B., Nurani, G., & Franzén, L.-G. 1996. Characterisation of the H^+-ATPase in plasma membranes isolated from the green alga *Chlamydomonas reinhardtii*. *Physiologia Plantarum* 97:445-453.

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* 22:3105-3117.

Pollock, S.V., Prout, D.L., Godfrey, A.C., Lemaire, S.D., & Moroney, J.V. 2004. The *Chlamydomonas reinhardtii* proteins Ccp1 and Ccp2 are required for long-term growth, but are not necessary for efficient photosynthesis, in a low-CO_2 environment. *Plant Mol Biol* 56:125-132.

Ramazanov, Z., Mason, C.B., Geraghty, A.M., Spalding, M.H., & Moroney, J.V. 1993. The Low CO_2-Inducible 36-Kilodalton Protein Is Localized to the Chloroplast Envelope of *Chlamydomonas reinhardtii*. *Plant Physiol* 101:1195-1199.

Raven, J.A. 2013. Rubisco: still the most abundant protein of Earth? *New Phytol* 198:1-3.

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.

Ricklefs, R.E., & Miller, G.L. 2000. Ecology. W. H. Freeman.

Rochaix, J.D. 2017. The Pyrenoid: An Overlooked Organelle Comes out of Age. *Cell* 171:28-29.

Schmidt, M., Gessner, G., Luff, M., Heiland, I., Wagner, V., Kaminski, M., Geimer, S., Eitzinger, N., Reissenweber, T., Voytsekh, O., Fiedler, M., Mittag, M., & Kreimer, G. 2006. Proteomic analysis of the eyespot of *Chlamydomonas reinhardtii* provides novel insights into its components and tactic movements. *Plant Cell* 18:1908-1930.

Somanchi, A., & Moroney, J.V. 1999. As *Chlamydomonas reinhardtii* acclimates to low-CO_2 conditions there is an increase in cyclophilin expression. *Plant Mol Biol* 40:1055-1062.

Spalding, M.H. 2009. Chapter 8 - The CO_2-Concentrating Mechanism and Carbon Assimilation A2 - Harris, Elizabeth H. Pp. 257-301 in: Stern, D.B., & Witman, G.B., (eds), *The Chlamydomonas Sourcebook (Second Edition)*. Academic Press, London. p 257-301.

Spalding, M.H., & Jeffrey, M. 1989. Membrane-Associated Polypeptides Induced in *Chlamydomonas* by Limiting CO_2 Concentrations. *Plant Physiol* 89:133-137.

Spalding, M.H., Spreitzer, R.J., & Ogren, W.L. 1983a. Carbonic Anhydrase-Deficient Mutant of *Chlamydomonas reinhardii* Requires Elevated Carbon Dioxide Concentration for Photoautotrophic Growth. *Plant Physiol* 73:268-272.

Spalding, M.H., Spreitzer, R.J., & Ogren, W.L. 1983b. Genetic and physiological analysis of the CO_2-concentrating system of *Chlamydomonas reinhardii*. *Planta* 159:261-266.

Spalding, M.H., Spreitzer, R.J., & Ogren, W.L. 1983c. Reduced Inorganic Carbon Transport in a CO_2-Requiring Mutant of *Chlamydomonas reinhardii*. *Plant Physiol* 73:273-276.

Sunderhaus, S., Dudkina, N.V., Jänsch, L., Klodmann, J., Heinemeyer, J., Perales, M., Zabaleta, E., Boekema, E.J., & Braun, H.P. 2006. Carbonic anhydrase subunits form a matrix-exposed domain attached to the membrane arm of mitochondrial complex I in plants. *J Biol Chem* 281:6482-6488.

Thompson, M.D., Mittelmeier, T.M., & Dieckmann, C.L. 2017. Chlamydomonas: The Eyespot. Pp. 257-281 in: Hippler, M., (ed), *Chlamydomonas: Molecular Genetics and Physiology*. Springer International Publishing, Cham. p 257-281.

Ueki, N., Ide, T., Mochiji, S., Kobayashi, Y., Tokutsu, R., Ohnishi, N., Yamaguchi, K., Shigenobu, S., Tanaka, K., Minagawa, J., Hisabori, T., Hirono, M., & Wakabayashi, K. 2016. Eyespot-dependent determination of the phototactic sign in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 113:5299-5304.

Umen, J.G., & Goodenough, U.W. 2001a. Chloroplast DNA methylation and inheritance in *Chlamydomonas*. *Genes Dev* 15:2585-2597.

Umen, J.G., & Goodenough, U.W. 2001b. Control of cell division by a retinoblastoma protein homolog in *Chlamydomonas*. *Genes Dev* 15:1652-1661.

Vahrenholz, C., Riemen, G., Pratje, E., Dujon, B., & Michaelis, G. 1993. Mitochondrial DNA of *Chlamydomonas reinhardtii*: the structure of the ends of the linear 15.8-kb genome suggests mechanisms for DNA replication. *Curr Genet* 24:241-247.

Van, K., & Spalding, M.H. 1999. Periplasmic carbonic anhydrase structural gene (*Cah1*) mutant in *Chlamydomonas reinhardtii*. *Plant Physiol* 120:757-764.

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.

Vance, P., & Spalding, M.H. 2005. Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Canadian Journal of Botany* 83:796-809.

Villarejo, A., Shutova, T., Moskvin, O., Forssén, M., Klimov, V.V., & Samuelsson, G. 2002. A photosystem II-associated carbonic anhydrase regulates the efficiency of photosynthetic oxygen evolution. *EMBO J* 21:1930-1938.

Wang, L., Yamano, T., Takane, S., Niikawa, Y., Toyokawa, C., Ozawa, S.I., 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 U S A* 113:12586-12591.

Wang, Y., & Spalding, M.H. 2006. An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 103:10110-10115.

Wang, Y., & Spalding, M.H. 2014a. 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.

Wang, Y., & Spalding, M.H. 2014b. LCIB in the *Chlamydomonas* CO_2-concentrating mechanism. *Photosynth Res* 121:185-192.

Wang, Y., Sun, Z., Horken, K.M., Im, C.-S., Xiang, Y., Grossman, A.R., & Weeks, D.P. 2005. Analyses of CIA5, the master regulator of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*, and its control of gene expression. *Canadian journal of botany* 83:765-779.

Whitney, S.M., Houtz, R.L., & Alonso, H. 2011. Advancing our understanding and capacity to engineer nature's CO_2-sequestering enzyme, Rubisco. *Plant Physiol* 155:27-35.

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.

Yamano, T., & Fukuzawa, H. 2009. Carbon-concentrating mechanism in a green alga, *Chlamydomonas reinhardtii*, revealed by transcriptome analyses. *J Basic Microbiol* 49:42-51.

Yamano, T., Miura, K., & Fukuzawa, H. 2008. Expression analysis of genes associated with the induction of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiol* 147:340-354.

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 U S A* 112:7315-7320.

Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S., 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.

Ynalvez, R.A., Xiao, Y., Ward, A.S., Cunnusamy, K., & Moroney, J.V. 2008. Identification and characterization of two closely related beta-carbonic anhydrases from *Chlamydomonas reinhardtii*. *Physiol Plant* 133:15-26.

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* 16:1466-1477.

Zelitch, I., Schultes, N.P., Peterson, R.B., Brown, P., & Brutnell, T.P. 2009. High glycolate oxidase activity is required for survival of maize in normal air. *Plant Physiol* 149:195-204.

CHAPTER 2. IDENTIFICATION OF THE *SU1* LOCUS

Introduction

Rubisco (ribulose-1,5-bisphosphate carboxylase/oxygenase) catalyzes the RuBP (ribulose-1,5-bisphosphate) carboxylation reaction – the principal CO_2-fixing step in photosynthesis – as well as the competing RuBP oxygenation reaction, which inhibits carboxylation and leads to the wasteful photorespiratory pathway. This enzyme has a tendency to pursue the latter pathway under current atmospheric CO_2 and O_2 concentrations and is notorious for its slow catalytic rate, so some terrestrial plants have evolved to employ CO_2-concentrating mechanisms (CCMs), such as the C4 and CAM (Crassulacean Acid Metabolism) pathways, that ensure assembly of higher concentrations of CO_2 around Rubisco, thus enabling CO_2 and the Rubisco carboxylation reaction to outcompete O_2 and the Rubisco oxygenase reaction.

Apart from the shortcomings of Rubisco, variability in ambient CO_2 levels is also a factor that these photosynthesizing organisms have to contend with, especially in aquatic environments. Aquatic photosynthetic organisms, such as the green alga *Chlamydomonas reinhardtii* (henceforth referred to as "Chlamydomonas"), are additionally disadvantaged because the rate of diffusion of CO_2 in water is about 10,000 times slower than it is in air (Evans and Von Caemmerer, 1996; Ricklefs and Miller, 2000). Algae, for example, have therefore developed their own inducible CCM, which is geared toward improving acquisition of inorganic carbon (commonly referred to as Ci) into their cells. In aquatic environments, Ci can be present in the forms of CO_2, HCO_3^-, and, more rarely, CO_3^{2-}, and the CO_2 solubility and the pH of the environment determine the equilibrium between these forms (Dodds and Whiles, 2010; Umen and Olson, 2012).

When induced, the CCM brings about a cascade of changes in gene expression, metabolism, and physiology, encompassing (i) induction/activation of one or more Ci uptake systems; (ii) carbonic anhydrases for interconversion between the different forms of Ci; and (iii) compartmentalization of Rubisco into a specialized structure (called a pyrenoid in algae, and a carboxysome in cyanobacteria) within the chloroplast where the enzyme is sequestered and where CO_2 can be delivered in high concentrations to encourage Rubisco carboxylase rather than oxygenase activity.

The Chlamydomonas CCM is induced in limiting CO_2 levels – low or air-level CO_2 (LCO_2; 0.03–0.4% CO_2) and very low CO_2 ($VLCO_2$; <0.02% CO_2) – as opposed to high CO_2 (HCO_2; >0.4% CO_2) levels (Vance and Spalding, 2005). In LCO_2, a limiting CO_2-inducible protein, LCIB, has been found to be necessary for survival, since *LCIB* mutants, *pmp1* and *ad1*, exhibit an LCO_2-lethal/*air dier* phenotype (Spalding *et al.*, 1983; Wang and Spalding, 2006). While the absence of LCIB does not cause lethality in $VLCO_2$, Wang and Spalding (2014a) found that it does still play a role: it is involved in active CO_2 uptake and works in concert with the active HCO_3^- uptake pathway involving putative HCO_3^- transporters, which include the plasma membrane-localized HLA3 and the chloroplast envelope-localized LCIA.

LCIB forms a 350 kDa hetero-multimeric complex with a homologous protein, LCIC, and, in HCO_2 and LCO_2, the complex is seen dispersed around the chloroplast stroma, while it is condensed around the boundaries of the pyrenoid in $VLCO_2$ (Wang and Spalding, 2014a, 2014b; Yamano *et al.*, 2010). In the absence of LCIB, active CO_2 uptake decreases dramatically, especially in LCO_2 (Wang and Spalding 2014a), and the cells are unable to retain Ci in air levels of CO_2 (Duanmu *et al.* 2009).

Epistatic interactions between mutations in *LCIB* and *CAH3* (a gene encoding a thylakoid carbonic anhydrase) revealed that LCIB functions downstream of CAH3 (Duanmu *et al.*, 2009), suggesting that LCIB may be involved in preventing CO_2 from leaking out of the cell. In addition, detailed photosynthetic interactions between *LCIB* and *LCIA* mutants revealed that LCIB is required for active CO_2 uptake (Wang and Spalding 2014a), in addition to preventing CO_2 leakage. Whether this is in fact what its function is, and if so, the molecular mechanism by which it does it, are yet to be definitively ascertained.

One possibility is that the LCIB/LCIC complex takes up externally-derived CO_2 and prevents CO_2 leakage by converting CO_2 into HCO_3^-, acting more-or-less as a unidirectional carbonic anhydrase (Wang and Spalding, 2014a). This possibility is supported by the structure of LCIB. Jin *et al.* (2016) found that LCIB and its homologs have structural similarities with β-carbonic anhydrases, and posited a theory that, in order to hydrate CO_2 (and therefore stop its escape from the chloroplast), the LCIB/LCIC complex would have to be regulated so that it functions only when there is more CO_2 than HCO_3^- in the chloroplast stroma, rather than working to maintain an equilibrium between the two Ci species at all times.

In order to gain more insight into the function of the LCIB/LCIC complex and its role and molecular mechanism in the CCM, Duanmu *et al.* (2011) performed insertional mutagenesis on *pmp1* and *ad1*, and screened for *air dier*-suppressors. They recovered eight independent suppressor lines, representing six different suppressor loci. Identifying the defective genes and encoded proteins responsible for such suppression could provide us insight into what LCIB itself does and how specific suppressor mutations allow for

growth in LCO_2 in the absence of a functional LCIB. One locus was identified as encoding CAH3 (Duanmu et al, 2009), but until now, none of the other suppressor loci have been identified.

One suppressor mutation, *su1*, was isolated in the *pmp1* background as strain *pmp-su1*. Unfortunately, this *su1* suppressor phenotype did not exhibit cosegregation with the paromomycin-resistance insert used for insertional mutagenesis, indicating the *su1* mutation arose through another mechanism, possibly related to the transformation process. Regardless of the molecular cause of the *su1* mutation, it was not tagged with the paromomycin-resistance insert, so identification of the *SU1* locus required alternative approaches.

In this project, we sought to identify the locus of and characterize the *su1* mutation. We employed an approach that involved crossing the mutant strain with a polymorphic, wild-type strain; screening and collection of independent progeny that demonstrated a non-mutant phenotype (and, therefore, also the non-mutant genotype in the locus of interest); and pooling of genomic DNA from each progeny, followed by sequencing of the pool. Sequenced genomes allowed for the creation of a map of single-nucleotide polymorphisms (SNPs) in the polymorphic strain relative to those in the database of references strains (which are a set of various wild-type strains; Gallaher *et al.*, 2015); this map was used to identify the region in the pooled DNA where the sequence was revealed to have SNPs only from the polymorphic parent. Once a broad region was identified, a gene-wise comparison was made between sequences of the pooled DNA, the mutant strain, and wild-type reference strains in order to shorten the list of candidate

mutant genes; one of these genes was then confirmed as being in the *SU1* locus by cosegregation of new mutants allelic to *su1* and the *air dier* suppressor phenotype.

Materials and Methods

Generation of Initial Progeny: Strains, Growth Conditions, and Crosses

Initial crosses for the generation of a recombinant progeny pool were between *pmp-su1* (*mt*$^+$; Duanmu and Spalding (2011)) and CC-2290 (also denoted S1 D2; *mt*$^-$; Gross *et al.* (1988)). Continuous gas flow was maintained to growth chambers with three different CO_2 levels: HCO_2 (5% [v/v] CO_2 enriched with air); LCO_2 (normal air, with 300–450 ppm CO_2); and $VLCO_2$ (normal air passed through saturated sodium hydroxide solution and then remixed with normal air). All chambers were kept at room temperature and under continuous illumination of 50–80 μE m^{-2} s^{-1}. Cells were grown and maintained on agar plates with TAP medium (Gorman and Levine, 1965) or with CO_2-minimal medium (Geraghty *et al.*, 1990) in a HCO_2 chamber. When testing growth in different CO_2 levels, only CO_2-minimal medium plates were used.

Crosses were performed based on the protocol described in Harris (1989): mating type *plus* and *minus* (*mt*$^+$ and *mt*$^-$) parental cells were resuspended in two separate 15 ml tubes with about 3 ml of nitrogen-free TAP liquid medium (nitrogen-free conditions are required to induce gametogenesis). After gentle shaking of the cells on an orbital platform shaker overnight in the light, about 0.5 ml of *mt*$^+$ and 0.5 ml of *mt*$^-$ cell suspensions were mixed and placed (at rest) in the light for 1-2 hours to encourage mating. Any pellicles that were formed were broken apart by shaking the tubes, following which about 0.5 ml of the mixture (of *mt*$^+$ and *mt*$^-$ cells) was spread on nitrogen-free TAP plates with 4% agar. The plates were placed first in the light overnight and then in

the dark (wrapped in aluminum foil) for 4-6 days; the initial exposure to light is important for zygospore formation and maturation. After the incubation in the dark, the vegetative, unmated cells were scraped off the medium using a sterilized razor blade; the high percentage of agar used (4%) makes it harder to gouge the medium in this step. This leaves behind the zygospores, which are embedded in the growth medium. The plates were then placed in a HCO_2 chamber in the light for 5-6 days to facilitate germination, (when the zygospore undergoes meiosis, producing four daughter zoospores) followed by the subsequent vegetative growth phase (reproduction of each zoospore by mitosis). At this point, a zygote colony will be visible with the naked eye on the growth medium.

Each zygote colony from multiple crosses was streaked on TAP plates for single colonies, ten of which were chosen at random and their growth phenotypes assessed after 5-7 days of growth on CO_2-minimal medium plates in different ambient CO_2 levels: HCO_2, LCO_2 (at about 300–450 ppm), and $VLCO_2$ (at about 100 ppm). Of the ten colonies randomly selected from each zygote colony, only one of any that exhibited the *air dier* phenotype was selected for further analysis. This ensured that no two selected progeny were siblings (i.e., from the same zygote). Each of 489 such progeny was spot-tested in the three different CO_2 levels to ensure that they had a lethal phenotype in air but grew in HCO_2 and $VLCO_2$. Spot tests were performed by suspending cells in liquid minimal medium to a density of about 10^6 cells/ml (cells counted using a hematocytometer from Reichert Scientific Instruments, Buffalo, NY; see Harris (1989) for protocol) and spotting equal volumes (usually 5μl or 10μl) on CO_2-minimal medium plates.

The progeny were screened *for* the *air dier* phenotype and *against* the suppressor phenotype to ensure they inherited the *LCIB* allele from the *pmp-su1* parent, and the *SU1* allele from S1 D2 (and, therefore, not from the *su1* mutation-carrying parent, *pmp-su1*); this eliminated the need for any additional work to confirm the genotypes of the two relevant loci. Screening for the *air dier* phenotype also ensured that the selected progeny were bona fide recombinant progeny (since neither of the parents had the *air dier* phenotype; see Figure 8 and Table 1) and not vegetative colonies from either parent used in the cross.

DNA Isolation and Measurement

To isolate genomic DNA, cells were grown on TAP or CO_2-minimal medium plates and harvested by resuspending them in 500 µl of DNA extraction buffer (100 mM Tris-Cl [pH 8.0], 20 mM EDTA, 500 mM NaCl, 2% sarkosyl) in a 1.5 ml centrifuge tube, at a concentration of about 2–5 x 10^6 cells/ml (all centrifugation steps were performed at room temperature and all tubes used were 1.5 ml Eppendorf centrifuge tubes). After vortex-mixing for about 3 minutes, 500 µl of equilibrated phenol (pH 8.0) was added in the fume hood and the suspension was vortexed vigorously for 60 minutes. Following centrifugation for 5 minutes, the upper aqueous layer was transferred to a new centrifuge tube, 500 µl chloroform was added, then vortex mixed and centrifuged for 5 minutes. The aqueous layer was transferred to a new tube, and, in order to digest RNA in the solution, 3 µl RNaseA (10mg/ml) was added and the tube incubated at 37°C for 30 minutes. To precipitate the DNA, the aqueous layer was transferred to a new tube, then 350 µl isopropanol (2-propanol) was added and mixed by inverting the tube several times. After

centrifugation at full speed for 18 minutes, the DNA pellet was washed with 500 µl of 70% ethanol, centrifuged at full speed for 5 minutes, decanted, and the pellet air-dried of ethanol. The DNA pellet was then resuspended in 40 µl of ddH_2O. A sample of the isolated DNA was electrophoresed through an agarose gel, and its concentration and purity were measured using a NanoDrop® ND-2000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). Genomic DNA was isolated from the 489 independent, recombinant, *air dier* progeny, and about 4 ng of DNA from each of them was pooled, with the final pool amounting to about 2 µg of DNA.

SNP Library, Sequencing of the Progeny DNA Pool, and Mapping of the Reads

A bar-coded library of the DNA pool from the 489 progeny was generated at the Iowa State University DNA Facility. Sequencing was performed at the University of California Los Angeles Broad Stem Cell Research Center (BSCRC) sequencing core on a HiSeq2000 sequencer (Illumina). The raw sequences were aligned to the *C. reinhardtii* reference sequence (strain CC-503) version 5 with BWA mem, version 0.7.5a-r405 (Li and Durbin, 2009) using default parameters. Duplicate read pairs were removed using Picard MarkDuplicates, version: 1.85(1345) (http://picard.sourceforge.net) with default parameters. The Genome Analysis Toolkit (GATK), version 2.6-5-gba531bd (DePristo *et al.*, 2011; McKenna *et al.*, 2010) was used to call variants on the aligned, de-duplicated reads.

SU1 Locus Identification by Deep Sequence Mapping

Using a previously constructed library of SNPs that occur across the S1 D2 nuclear genome (Gallaher *et al.*, 2015), the sequence of the pooled DNA of the progeny was characterized based on the abundance of SNPs in it.

A ratio was calculated for every position with a SNP call in the progeny DNA pool. A base position was said to have a "SNP call" if at least one of all sequencing reads of that position in the progeny DNA pool did not match the base in the same position in the sequence from the database of reference strains, i.e. if a non-WT (non-wild-type) SNP was observed in the pool DNA sequence.

Thus, this ratio was the number of non-WT SNPs encountered as a fraction of all the reads:

$$\frac{\text{\# of times a } \boxed{\textbf{non-WT}} \text{ SNP is read}}{\boxed{\textbf{total}} \text{ \# of reads}} \longrightarrow \text{Formula I}$$

The ratios can be understood based on the *LCIB* and *SU1* genotypes of the progeny and on how they are calculated in three types of regions of the genome:

1. at the *SU1* locus
2. at the *LCIB* locus
3. in the regions of the genome other than the *LCIB* and *SU1* loci

The following explanations for how the ratios were calculated assume that the sequencing coverage was 50X:

I. At the *SU1* locus in the pooled DNA:

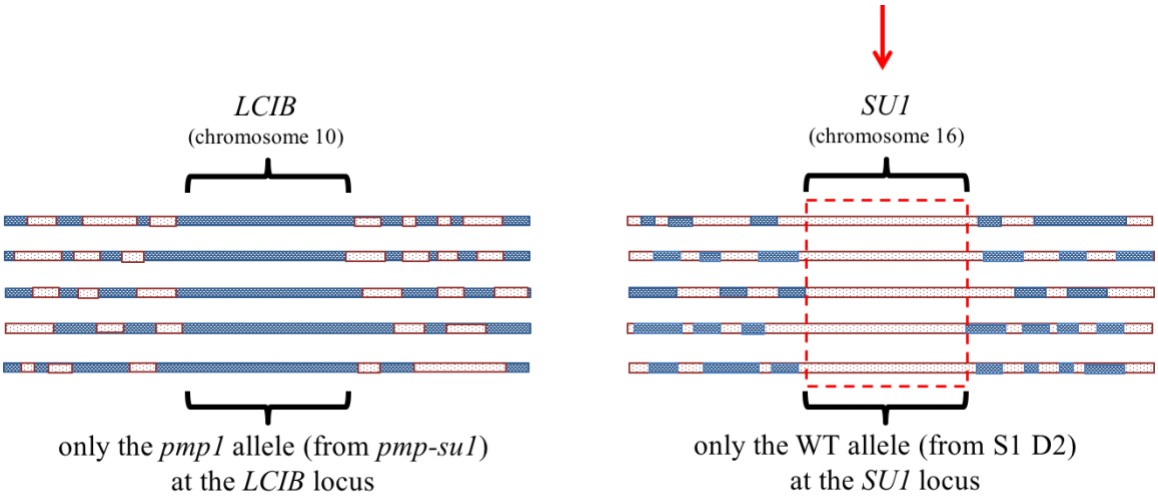


Figure 3. The different regions of the genome in five example progeny collected, with the red ( ) region representing genome inherited from the S1 D2 parent, and the blue ( ) region representing genome from the *pmp-su1* parent (see Figure 8). Here, the *SU1* locus is highlighted with a red arrow.

If there was an A > T SNP change between the reference sequence and S1 D2, ideally, all 50 out of the 50 reads should be T at that particular base position, since all the progeny were isolated so that the DNA at this locus is inherited ONLY from the S1 D2 parent (Figure 3).

The sequence in the pool can be annotated as – – – – – A – – – – – x 0

– – – – – T – – – – – x 50

The calculated ratio would be $\frac{50}{50} = 1$

- The 50 in the numerator represents the number of times a non-WT SNP (ideally, always) was read in the pooled DNA sequence
- The 50 in the denominator represents the total number of reads

II. At the *LCIB* locus in the pooled DNA:

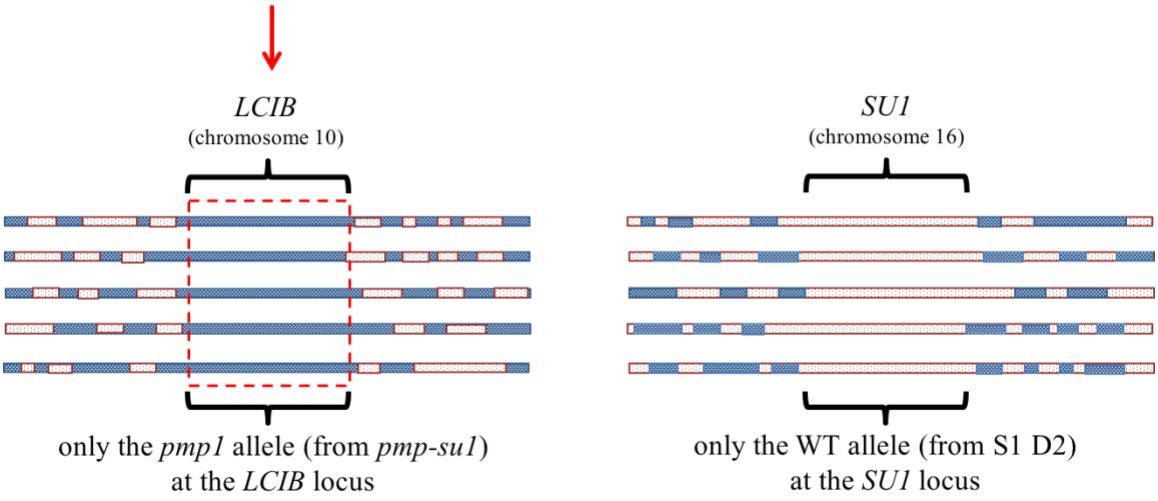


Figure 4. The different regions of the genome in five example progeny collected, with the red ( ) region representing genome inherited from the S1 D2 parent, and the blue ( ) region representing genome from the *pmp-su1* parent (see Figure 8). Here, the *LCIB* locus is highlighted with a red arrow.

If there was an A > T SNP change between the reference sequence and S1 D2, ideally, NONE of the 50 reads should be T, since all the progeny were isolated so that the DNA at the *LCIB* locus (here, the *pmp1* allele) is inherited ONLY from the *pmp-su1* parent (Figure 4).

The sequence in the pool can be annotated as

– – – – – A – – – – – x 50

– – – – – T – – – – – x 0

The calculated ratio would be $\frac{0}{50} = 0$

- The 0 in the numerator represents the number of times a non-WT SNP (ideally, never) was read in the pooled DNA sequence
- The 50 in the denominator represents the total number of reads

III. In the regions of the genome other than the *LCIB* and *SU1* loci:

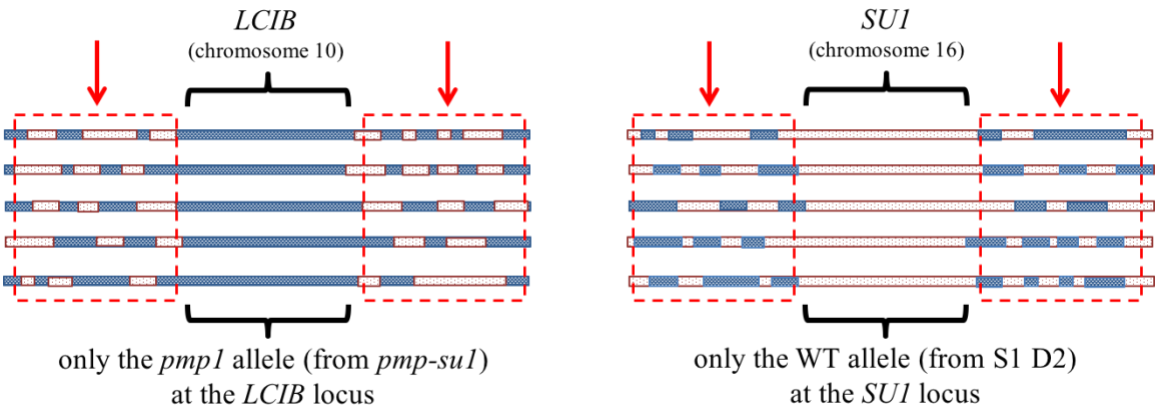


Figure 5. The different regions of the genome in five example progeny collected, with the red ( ) region representing genome inherited from the S1 D2 parent, and the blue ( ) region representing genome from the *pmp-su1* parent (see Figure 8). Here, the non-*LCIB*, non-*SU1* loci are highlighted with red arrows.

If there was an A > T SNP between the reference sequence and S1 D2, some reads will be T (meaning the DNA was inherited from the S1 D2 parent) and some will be A (meaning the DNA was inherited from the *pmp-su1* parent; see Figure 5).

Considering an example in which 30 of 50 reads were T and the other 20 reads were A:

The sequence in the pool can be annotated as – – – – – A – – – – – x 20

– – – – – T – – – – – x 30

The calculated ratio would be $\frac{30}{50} = 0.6$

- The 30 in the numerator represents the number of times a non-WT SNP was read in the pooled DNA sequence
- The 50 in the denominator represents the total number of reads

This ratio would be some number between 0 and 1, depending on how many reads from the pooled DNA sequences at a particular nucleotide position were non-WT SNPs and how many were WT SNPs.

The expected ratios are:

- 1, at the *SU1* locus
- 0, at the *LCIB* locus
- (0, 1), at every other locus

Artificial MicroRNA Knockdown of *LCI15*, the Putative *SU1* Allele

The Web MicroRNA Designer platform (WMD3, http://wmd3.weigelworld.org/cgi-bin/webapp.cgi) was used to identify suitable artificial microRNA (amiRNA) candidates to target mRNA from *LCI15* with minimal off-target effects. Two 90-bp oligonucleotides, amiL15F1 (5'–CTAGTCTGTCCCTCTATGTCGGTAAATCTCGCTG ATCGGCACCATGGGGGTGGTGGTGATCAGCGCTATTTATTGACATAGAGGGA CAGG–3') and amiL15F2 (5'–CTAGCCTGTCCCTCTATGTCAATAAATAGCGCTG ATCACCACCACCCCCATGGTGCCGATCAGCGAGATTTACCGACATAGAGGGA CAGA–3'), that had sequences targeting exon 5 in *LCI15* (+1115 to + 1135 bps downstream of the start codon), were annealed and the resulting dsDNA was ligated into the *pChlamiRNA3* vector using the *SpeI* resistriction enzyme site for the ligation, as described in Molnar *et al.* (2009). *pChlamiRNA3* has the *AphVIII* gene (Sizova *et al.*, 2001), under the control of the constitutive *HSP70A-RBCS2* promoter and *RBCS2* terminator, as a selectable marker. The amiRNA precursor was under the control of a

PSAD promoter and terminator (Fischer and Rochaix, 2001; Molnar *et al.*, 2009). The resulting vector was transformed into the strain *pmp1* as described below. Transformants were selected on TAP (tris-acetate-phosphate) agar plates supplemented with 15 mg/L of paromomycin and screened on CO_2-minimal medium plates in the three CO_2 levels (Figure 17) as described above.

Transformation Procedure

The protocol in Shimogawara *et al.* (1998) was optimized by Drs. Yingjun Wang and David Wright as follows: for each transformation, 25 ml of culture was grown in TAP (tris-acetate-phosphate) medium to a concentration of 0.4–1 x 10^6 cells/ml and harvested by centrifuging at 1000 *g* for 5-10 minutes. About 2.5 µg of plasmid DNA was linearized by overnight digestion with an appropriate restriction enzyme, small amounts of which were added in 3 or 4 installments. After the digestion, the enzyme was heat inactivated and the DNA was filter sterilized using Corning® Costar® Spin-X® centrifuge tube filters (0.22 µm, sterile; catalog # *CLS8160*, Sigma-Aldrich, Inc.). A 4 mm gap electroporation cuvette was chilled at 16 ºC for at least 15 minutes. The supernatant of the centrifuged culture was removed, and the cells were resuspended in TAP medium supplemented with 0.06 M sucrose (0.25 ml of TAP+sucrose for 25 ml of original growth culture) and placed on ice for 10 minutes. The digested DNA (2.5 µg) was then mixed with the cell culture (0.25 ml) and the mixture was transferred to the chilled electroporation cuvette. The cuvette was then placed in 16 ºC for another 5 minutes. The electroporation was carried out at 650V, 25 µF, and 0 resistance in a Gene Pulser Xcell™ Electroporation system. The cells were rested at room temperature for 10 minutes and then resuspended in 50 ml TAP+sucrose medium. After shaking at 25 ºC for at least 24

hours, the culture was centrifuged at 1000 *g* for 5-10 minutes, the supernatant removed, the cell pellet resuspended in about 0.1 ml TAP+sucrose, and the resuspension spread on TAP plates supplemented with the appropriate antibiotic. The plates were incubated at 25 ºC in HCO_2 with light until colonies appeared (usually 7-10 days).

Complementation of the Putative *su1* Allele

The various vectors used for attempted complementation of *LCI15* mutant strains with a wild-type copy of *LCI15* genomic DNA or cDNA are listed in Table 4. In order to clone *LCI15* cDNA, CC-620 (wild type) cells were grown to mid-log phase in liquid minimal medium, bubbled with air-level CO_2 (350–400 ppm CO_2), and harvested. RNA was extracted using the RNeasy® Plant Mini Kit (Qiagen, Inc.) and cleaned of any contaminating DNA with the Ambion® TURBO DNA-free™ kit (Life Technologies Corporation).

A cDNA library was generated using the Invitrogen SuperScript® III First-Strand Synthesis System (Life Technologies Corporation) using about 50 ng of RNA as measured on a NanoDrop® ND-2000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). PCR using *LCI15*-specific primers (with appropriate restriction enzyme recognition sites incorporated) was performed to amplify cDNA from the gene using GoTaq® Green Master Mix (Promega Corporation). The size of the PCR products was verified by agarose gel electrophoresis, and the products were purified using the QIAquick® Gel Extraction Kit (Qiagen, Inc.) and sequenced (at the Iowa State University DNA Facility) to check for errors. Following this, the DNA fragments were digested with

appropriate restriction enzymes and cloned in to an expression vector (see Table 4 for list of vectors). Transformation was performed as described above.

Verification of *su1* as an Allele of *LCI15*: Strains, Crosses, and T7 Assays

Subsequent crosses used to verify *su1* as an allele of *LCI15* were performed between *ad1*, a *pmp1* allelic mutant (Wang and Spalding, 2006), and a *LCIB-LCI15* double mutant generated by employing a CRISPR/Cas9 system with a sgRNA targeting *LCI15* in a *pmp1* background. The novel CRISPR/Cas9 system, conceived and designed by Jiang and Weeks (2017), uses a gene-within-a-gene construct in which an artificial intron, with the sgRNA targeting the gene of interest inserted in it, is placed within the Cas9 gene. This consolidates the conventional construct with two separate Cas9 and sgRNA genes into a single gene. A 21-bp sequence (5'–GCGTGTACTTGTCAATCGCGG–3'; +575 to +595 downstream of the start codon) in exon 2 of *LCI15* was chosen as the target and the corresponding sgRNA inserted into the intron was 5'–GCGTGTACTTGTCAATCG–3'. The plasmid, with the Cas9/intron-sgRNA gene under the control of a *PSAD* promoter and terminator, was transformed into *pmp1* cells by electroporation as described above.

Based on the premise that *SU1* was an allele of *LCI15*, loss-of-function mutations in *LCI15* in an *LCIB*-mutant background would result in a suppressor phenotype which is easy to identify in a screen; this strategy was particularly advantageous when using this system, which lacks a selectable marker (such as an antibiotic resistance gene), to identify successful transformation events. Post-transformation, cells were plated on minimal medium plates, which were placed in LCO_2 chambers to allow for growth of *air dier*-suppressors. Twenty-five colonies that appeared on the plates were chosen, and a

681-bp region of their *LCI15* genomic sequence (encompassing the intended target in exon 2) was amplified by PCR (using primers L15Gn13FP 5'–CTGGAGCTGGCTTCGCCG–3' and L15Gn18RP 5'–GGATGTTGACCTAGTGCCAGACACTG–3') and sequenced to reveal any edits. Three colonies had both the suppressor phenotype *and* edits in their *LCI15* genomic sequence: Cas9-19 had a 2-bp deletion (+487-488 downstream of the start codon); Cas9-27 had a 7-bp deletion (+522 to +528 downstream of the start codon), and Cas9-28 had a 1-bp deletion (+548 from the start codon). All of these edits resulted in premature stop codons and truncated predicted LCI15 proteins (Figure 11).

Cas9-27 was chosen for further analysis and crosses, since the 7-bp deletion in its *LCI15* gene could be easily detected using T7 endonuclease assays. Crosses between *ad1* and Cas9-27 were carried out as described above, based on the protocol in Harris (1989). Phenotypic analysis at different CO_2 levels was performed as described above. To check for zeocin resistance of the progeny, minimal medium plates were supplemented with 10 μg/ml zeocin. Genotypic analysis to confirm the *LCI15* allele in the progeny were performed using T7 endonuclease assays as described below.

A 681-bp region of the *LCI15* genomic sequence, from each progeny and from a WT strain, was amplified by PCR (using primers L15Gn13FP and L15Gn18RP). A mixture of 7 μl of PCR product from one progeny, 7 μl of PCR product from the WT strain, and 13 μl of H_2O was heated and gradually cooled to denature and re-anneal the DNA sequences, thereby creating a mixture of hybrid dsDNA sequences (with one strand from the progeny DNA sequence and the other from the WT sequence) and reannealed, non-hybrid dsDNA sequences. To 10 μl of this mixture was added 0.5 μl of T7

endonuclease (New England Biolabs, Inc., #M0302) and the mixture was then incubated at 37°C for 2 hours. This T7 endonuclease-treated DNA was then electrophoresed through an agarose gel to check for cleavage of hybrid DNA sequences.

Based on the presence or absence of the 7-bp *LCI15* deletion in the progeny (Figure 13), determination of the *LCIB* allele as *ad1* (zeocin-resistant; see Figure 12) or *pmp1* (zeocin-sensitive), and determination of the mating type, progeny were verified as definitely bona fide, i.e., colonies that arose from a zygote in the cross, or as possibly arising from a parental vegetative cell (Table 2).

Four *air dier* suppressor mutants with inserts in *LCI15* were generated by insertional mutagenesis using a vector with the paromomycin resistance-conferring gene *AphVIII* using the strain *ad1* (Wang and Spalding, 2006). Two of these mutants, A45 and B28, were determined by PCR to have an insertion in the 3' half of *LCI15*, while the other two, B4 and B24, have an insertion in the 5' half of the gene. The transgene insertions rendered the respective gene halves unable to be amplified by PCR. The primers used were designed to specifically amplify either the 5' half of the *LCI15* gene (-427 to +1187 downstream of the start codon; the primers used were lci15-5a [5'–CCCGTACACCACAGGCTTGTC–3'] and lci15-5as [5'–GAAGCGGTACTCGTCCAAGG–3']), or the 3' half of the gene (+984 to +2599 downstream of the start codon; the primers used were lci15-3a [5'–GCCAGTGTCTGGCACTAGGT–3'] and lci15-3as [5'–TTCCTGCAGTGCTGCTATCA–3']).

The nature and components of the insertions were determined in two mutants, A45 and B4, by PCR:

- In A45, the promoter to terminator sequence of the *AphVIII* cassette from the pSI103 plasmid was inserted in exon 3 of *LCI15*, thus disrupting the coding sequence of the gene (Figure 6).

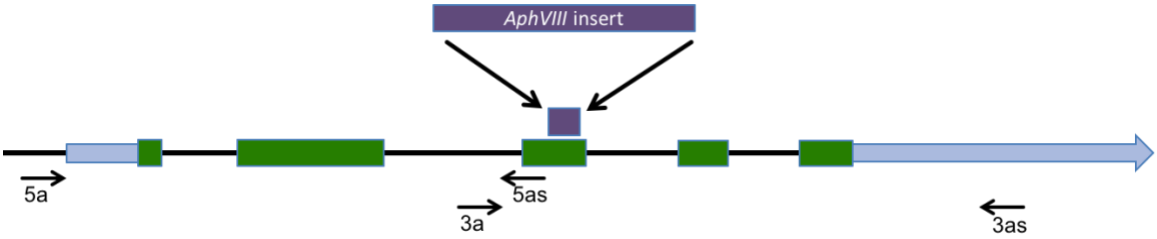


Figure 6. The *AphVIII* insertion in the *LCI15* gene in the mutant strain A45. The green boxes (▬) represent the exons of the endogenous gene. The purple box (▬) above exon 3 is the site of the insertion. The annealing sites of the primers used are also shown.

- In B4, part of the *AphVIII* cassette was inserted in intron 2 of *LCI15*, possibly disrupting the 3' end of exon 2 (Figure 7).

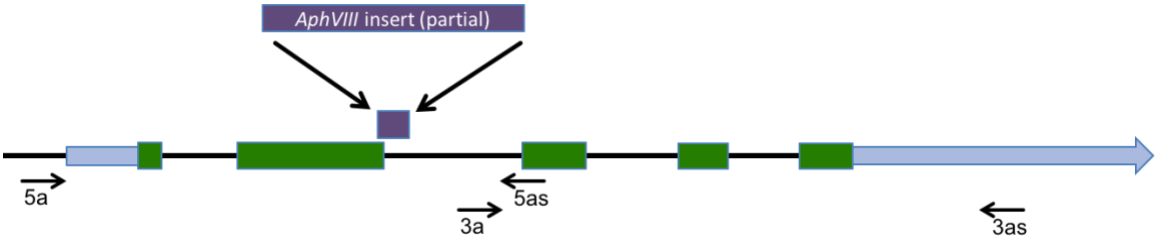


Figure 7. The *AphVIII* insertion in the *LCI15* gene in the mutant strain B4. The green boxes (▬) represent the exons of the endogenous gene. The purple box above (▬) exon 2 and intron 2 is the site of the insertion. The annealing sites of the primers used are also shown.

The insertions in the other two insertional mutants, B4 and B24, were not characterized. All four mutants are zeocin-resistant (*ad1* allele carries Ble^R gene in the *LCIB* locus) and paromomycin-resistant (by the $AphVIII/Par^R$ gene insertion).

Crosses were performed between the *LCIB* mutant *pmp1* (Spalding *et al.*, 1983) and each of the two characterized mutants mentioned here, A45 and B4. Phenotypic analysis at different CO_2 levels was performed as described above. To check for antibiotic resistance of the progeny, minimal medium plates were supplemented with either 10 μg/ml zeocin or 15 μg/ml paromomycin; zeocin to check for the presence of the *ad1 LCIB* allele, and paromomycin to check for the presence of the insertional mutant *LCI15* allele.

Results

Recombinant Progeny for Deep Sequence Mapping of *su1*

Strain CC-2290 (also called S1 D2) was crossed to *pmp-su1* to facilitate mapping of *su1* (Figure 8). S1 D2 has widespread single nucleotide polymorphism (SNP) base substitutions in comparison to the genome of commonly-used lab strains 21GR and 137c; SNPs in S1 D2 occur at an average frequency of about 1 in every 47 base pairs (bps), or 2.7 per 100 bps (Gross *et al.*, 1988; Kathir *et al.*, 2003; Vysotskaia *et al.*, 2001).

Using only independent (non-sibling), bona fide recombinant progeny (as explained in Materials and Methods) was critical for minimizing the number of progeny needed for mapping. To ascertain independence of the progeny from one another, only one progeny that arose from a zygote was included in the final progeny pool. Since all the progeny that were screened for were of the *air dier* phenotype, they were guaranteed to

be recombinant, since neither parent (*pmp-su1* and S1 D2) exhibited that phenotype and it would have had to arise from a recombination event (see Figure 9 and Table 1).

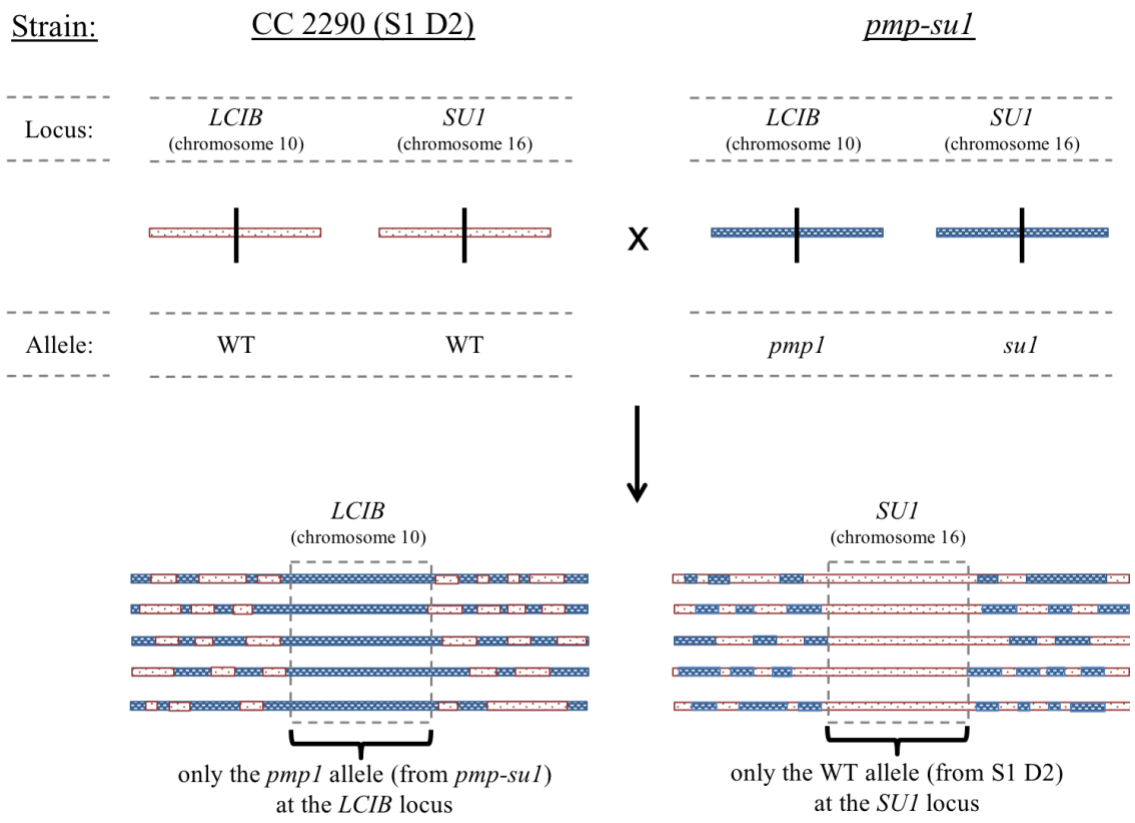


Figure 8. A schematic depiction of the initial cross that was carried out in this experiment. *pmp-su1* is the double-mutant parent which carries the *air dier*, *LCIB* mutation *pmp1* and its suppressor mutation, *su1*. The other parent, CC-2290 (also called S1 D2), is a polymorphic strain that has 1 SNP approximately every 47 bps. Only progeny that had the *pmp1* mutation but not the *su1* mutation were collected for the analysis. The red ( ) region represents DNA from S1 D2, and the blue ( ) region represents DNA from *pmp-su1*.

DNA from 489 independent, recombinant progeny was pooled, sequenced and aligned with the reference genome and a previously constructed library of S1 D2 SNPs as described in the Materials and Methods section. This allowed for ratios of non-WT or

non-reference SNPs as a fraction of all reads to be calculated for each SNP call in the progeny DNA pool.

Table 1. Possible genotypes and phenotypes of progeny in the *pmp-su1* x S1 D2 cross. The genotype and phenotype screened for and included in the pool of 489 progeny are highlighted in red.

Genotype of the progeny at the…		Phenotype of the progeny
LCIB locus	*SU1* locus	
pmp1	*su1*	WT
WT	WT	WT
pmp1	WT	*air dier*
WT	*su1*	WT

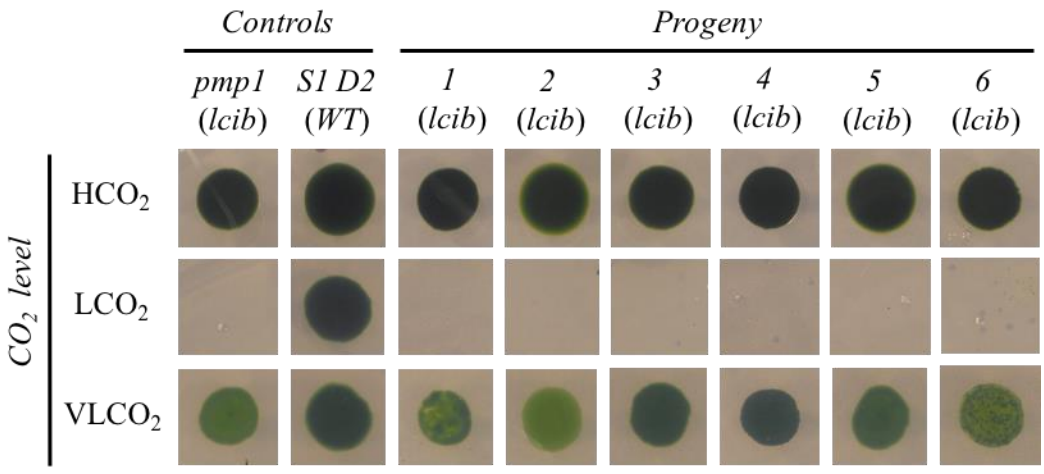


Figure 9. Spot tests for photoautotrophic growth in various CO_2 concentrations of progeny collected in the *pmp-su1* x S1 D2 cross, with the parents used as controls.

The average of all the SNP call ratios in each 10,000-bp region was calculated and plotted in Figure 10. This figure shows a "crest" that peaked at about 5,000,000 bps on chromosome 16 – presumably in the vicinity of the *SU1* locus, since the values of the ratios calculated at this locus were almost equal to 1. There was also a "trough" at the *LCIB* locus which is on chromosome 10; this functioned as a control for the experiment, because, since all progeny sequenced were of the *air dier* phenotype and thus the *LCIB* mutant genotype, the ratios *had* to be 0 or close to 0 at that gene locus.

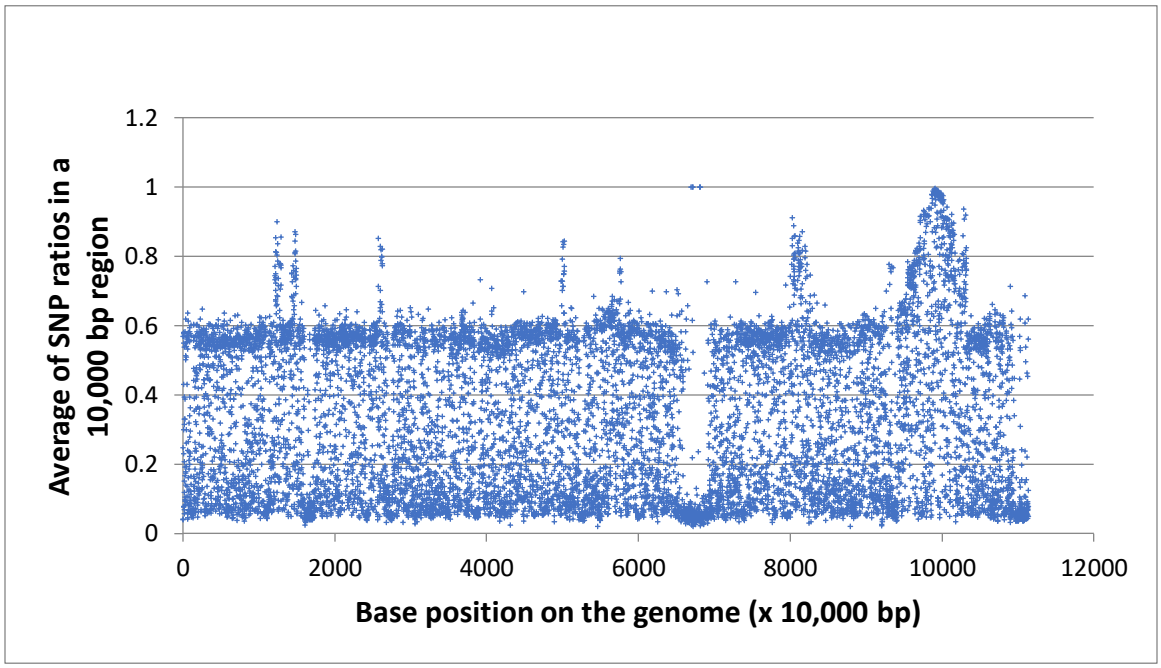


Figure 10. The graph above is a plot of the ratios calculated for each base position in the pooled DNA of the progeny (as described in Formula I) and then averaged over a 10,000-bp region *on the Y-axis* versus a scaled version of the entire genome *on the X-axis* (1 unit on the X-axis represents 10,000 bps of the genome). As seen between 6,000 and 8,000 units on the X-axis (chromosome 10), at the *LCIB* gene locus, the ratio values were calculated as close to or equal to 0. Roughly around 10,000 units on the X-axis (chromosome 16), the ratio values all were calculated as exactly or close to 1, which is the predicted average at the *SU1* locus.

One might naively expect a ratio of about 0.4 - 0.6 in regions other than near the *SU1* and *LCIB* loci. The observation that the actual ratios were spread fairly broadly between 0 and 0.6 in regions other than these two specific loci is easily explained. In order to avoid missing critically important diagnostic SNPs, the rules for SNP calls were not very stringent (see the Materials and Methods section), so false positives were likely quite common in the reads, i.e., what counted as a SNP could have actually been a sequencing error, and thus could have affected the SNP ratio calculated. For example, if out of 50 reads, all but one read matched that of the wild-type reference sequence (i.e., one of 50 reads was a non-WT SNP), it would still be counted as a SNP and the ratio for that base position would be 1/50 = 0.02. This will lower the overall average of the SNP ratios that fall within that 10,000 bp window. Such false positives were likely common in the SNP calling, which was fairly liberal.

Since ratios calculated at or near the *LCIB* locus were already 0 or almost 0, false positive calls' ratios would have little or no affect in those regions' overall ratios. In the *SU1* region and the non-*LCIB*, non-*SU1* regions of the genome, however, larger non-WT SNP reads-to-total reads fractions are generated, so a false positive-induced low ratio would affect them more adversely, lowering them.

As steps toward identifying potential gene candidates in the identified crest region, gene-wise SNP ratios were calculated for about a 7.7 million-bp region on chromosome 16, near the peak of the crest seen in Figure 10, and *pmp-su1*, one of the parent strains, was also sequenced and that sequence mapped to the reference strains. Based on near-1.0 SNP ratios, a 1,500,000-bp region of chromosome 16 (between bp positions 4,500,000 and 6,000,000) was selected as the most probable location for the

SU1 locus. The following rules were applied to map the *SU1* region more "finely" and the positions of interest were defined as those where:

1. **The *pmp-su1* sequence differs from the reference:** a novel mutation is expected at the *su1* locus, so the nucleotide sequence cannot be the same as that in the reference.
2. **The *pmp-su1* sequence differs from sequences of the pool and of S1 D2:** progeny selection for the DNA pool was based on the criterion that S1 D2 does not share the same phenotype as *pmp-su1*.
3. **The sequence change/mutation results in a change in predicted amino acid sequence**: changes in intron sequence (unless affecting splicing) and silent nucleotide changes in the coding sequence are unlikely to affect phenotype.

Identified sequence differences in two genes satisfied all three criteria:

1. *LCI15* (gene designation in the JGI Phytozome 12 database: Cre16.g685050)
    - G > T SNP change/mutation in the second base of exon 4 of the gene;
    - gGc > gTc (glycine to valine) change in the amino acid sequence
    - Low-CO_2 inducible gene, regulated by *CIA5/CCM1* (master regulator in the CCM)
    - Annotated as a PRL1 interacting factor L

2. Cre16.g682750
    - G > A SNP in intron
    - No annotated domains or functions

The SNP change in Cre16.g682750 is in intron 6 of the gene, but not in the 5'–GU or 3'–AG splicing consensus sequences. It is also unlikely that it would affect (or replace) the branch point base A, which, usually, is followed by a polypyrimidine tract and is within 20-50 bps upstream of the 3'–AG. So, although it was a candidate gene because there was a clear SNP change in *pmp-su1* compared to the reference, the change does not appear to have an obvious effect on predicted splicing consensus sequences.

LCI15, on the other hand, harbors a mutation in its coding sequence *and* is a LCO_2-inducible gene, clearly affected by the CCM. This made *LCI15* very attractive as a candidate gene (and Cre16.g682750 less so) and, therefore, we decided to pursue it first.

Verification of *su1* as an *LCI15* Allele

Multiple methods were employed to determine whether *LCI15* is, in fact, the gene mutated in the *su1* suppressor lines. Various new *LCI15* mutants were isolated in *LCIB* mutant backgrounds, and crosses between these and other *air dier* background strains were performed to determine whether cosegregation of the *LCI15* mutation and *air dier*-suppressor phenotype occurred. Other approaches, such as specific knock-down of *LCI15* expression in *pmp1* using artificial miRNAs, and complementation of the *su1* mutation in *pmp-su1* with a wild-type *LCI15* allele were also attempted.

Crosses between *LCIB-LCI15* double mutants and *LCIB* mutants (each parent with different *LCIB* mutant alleles)

Cas9-27, an LCI15 null mutant in an LCIB mutant background

A predicted null mutant in *LCI15* generated in a *pmp1* background during experiments employing a CRISPR/Cas9 genome-editing protocol (Jiang and Weeks, 2017) has a 7-bp deletion in exon 2 that leads to a frameshift mutation, a premature stop

codon, and thus a heavily truncated protein with 142 residues (WT LCI15 has 316 residues; see Figure 11).

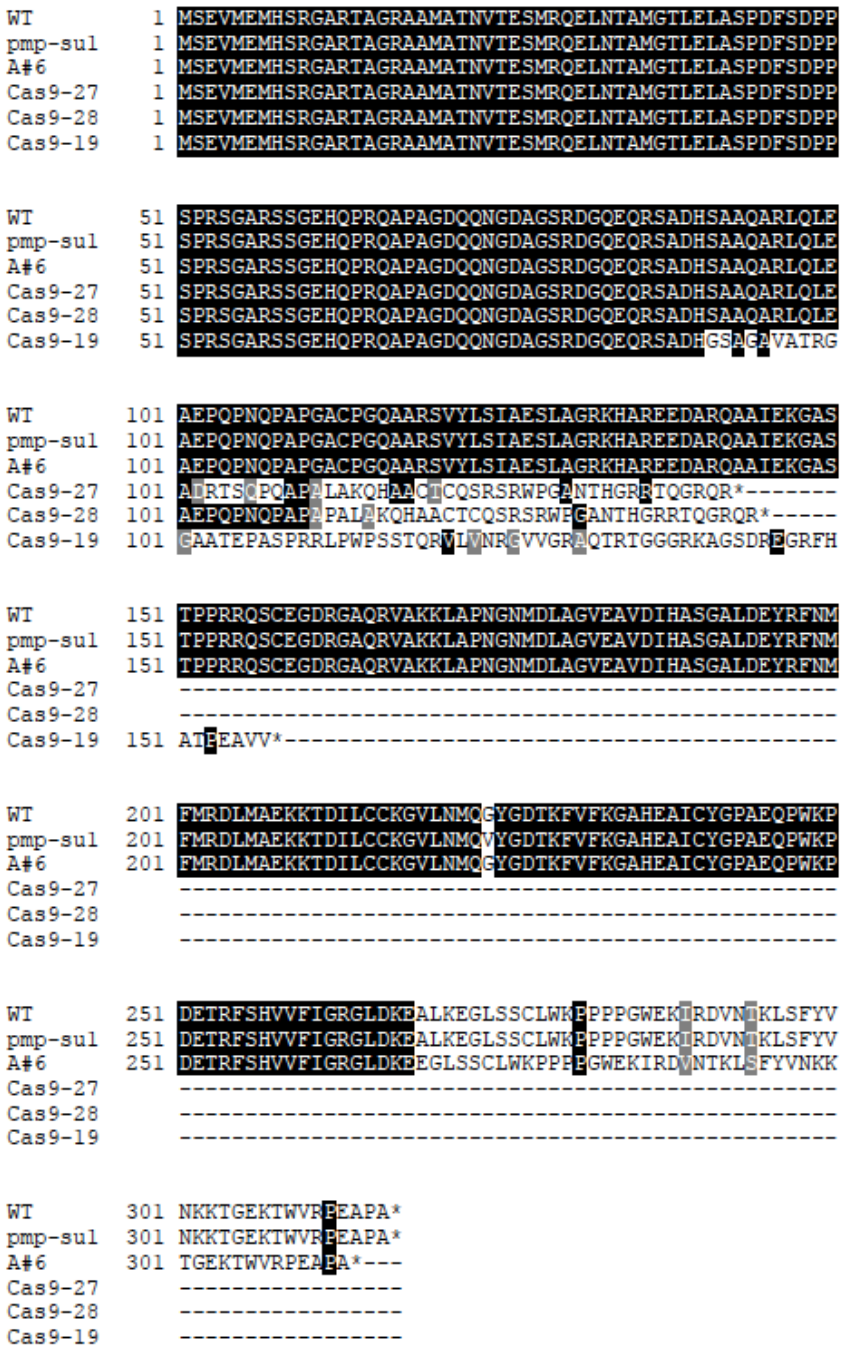


Figure 11. Sequence comparison of the predicted LCI15 protein product from wild-type and various mutant *LCI15* alleles.

Spot tests demonstrated that Cas9-27 suppresses the *air dier* phenotype of the *LCIB* mutant. Two other mutants, Cas9-28 and Cas9-19, which had 1-bp and 2-bp deletions in exon 2, respectively, were also isolated in the same CRISPR/Cas9 experiments and were confirmed to suppress the *air dier* phenotype of the *LCIB* mutant. These deletions similarly led to premature stop codons and truncated gene products (144 and 157 amino acids, respectively; see Figure 11).

It should be noted that the deletions in these three mutants did not occur within the CRISPR/Cas9 guide RNA's 21-bp target region, so we cannot confirm that the mutations resulted directly from CRISPR/Cas9 gene-editing itself, although all 3 deletions are within 100 bps upstream of the intended target, suggesting that there might have been some off-target editing that fortuitously resulted in mutations in our gene of interest, which were then identified during selection for putative editing events.

Table 2. Possible genotypes and phenotypes of progeny of the cross between Cas9-27 and *ad1*. For ease of reference, the two parental genotypes are labeled P1 and P2, and the two recombinant genotypes are labeled P3 and P4.

LCIB genotype	*LCI15* genotype	Zeocin resistant (+) / sensitive (–)	Suppressor (+)/ *air dier* (–)	T7 assay: cleavage (+) / no cleavage (–)	Progeny type
pmp1	Δ7bp	–	+	+	P1
ad1	WT	+	–	–	P2
ad1	Δ7bp	+	+	+	P3
pmp1	WT	–	–	–	P4

Cas9-27, which is in a *pmp1* background, was crossed to another *LCIB* mutant allele, *ad1* (Wang and Spalding, 2006), to confirm that the *LCI15* mutation in Cas9-27 suppresses the *air dier* phenotype of the other *LCIB* mutant allele and cosegregates with the suppression phenotype (Table 2).

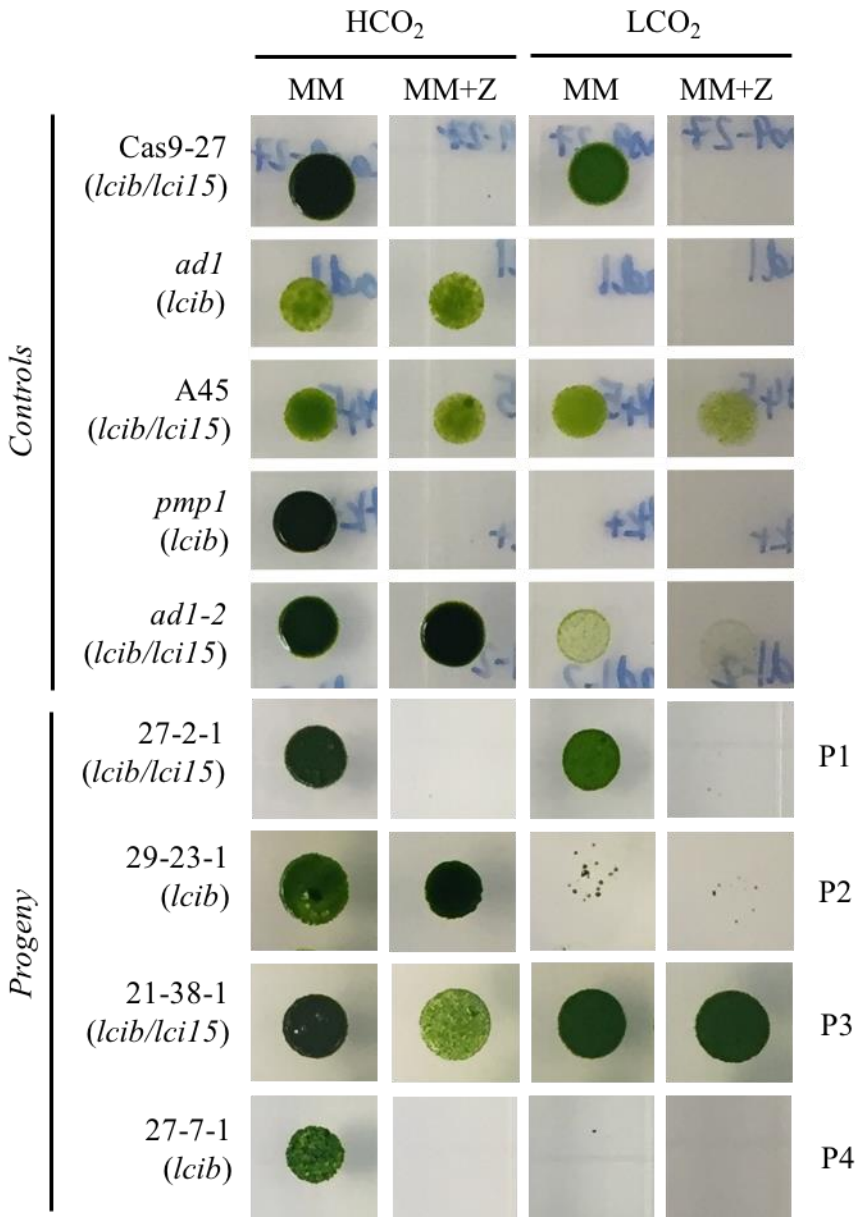
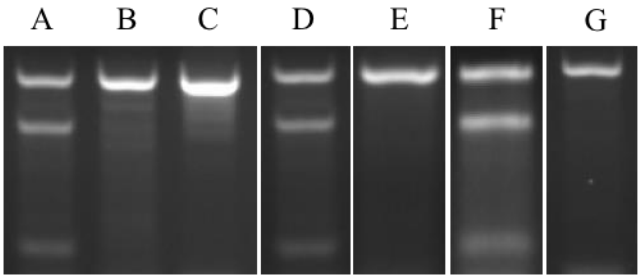


Figure 12. Spot tests of example progeny of the cross between Cas9-27 and *ad1*: one progeny of each of the four possible genotypes (P1–P4) of progeny are shown here. MM: minimal medium. MM+Z: minimal medium supplemented with zeocin.

Of the 33 bona fide progeny from this cross, all 12 that inherited the 7-bp deletion from Cas9-27 (detected by T7 assays; see Materials and Methods, and Figure 13) showed the suppressor phenotype, and the remaining 21 that inherited the wild-type *LCI15* allele retained the *air dier* phenotype (Figure 12).



DNA used as templates in PCRs for the T7 assay:

A:	[Cas9-27 + *ad1*]	(positive control)
B:	[*ad1* + *ad1*]	(negative control)
C:	[Cas9-27 + *ad1*], no T7	(no-enzyme negative control)
D:	[Progeny 27-2-1 + *ad1*]	(example of progeny type P1)
E:	[Progeny 29-23-1 + *ad1*]	(example of progeny type P2)
F:	[Progeny 21-38-1 + *ad1*]	(example of progeny type P3)
G:	[Progeny 27-7-1 + *ad1*]	(example of progeny type P4)

Figure 13. Gel electrophoresis of T7 assays of part of the *LCI15* genomic sequence of progeny from Cas9-27 and *ad1*.

Insertional mutants in LCIB mutant backgrounds

The four insertional mutants (described in Materials and Methods), all in an *ad1* background, were crossed to *pmp1* to verify whether the *LCI15* insertion in each is responsible for the suppressor phenotype, and to determine whether paromomycin resistance (from the *AphVIII* insert) cosegregates with the suppressor phenotype (Table 3).

Table 3. Possible genotypes and phenotypes of progeny of the cross between different *LCIB-LCI15* double mutants and *pmp1*. The four double mutants used were A45, B4, B24, and B28. For ease of reference, the two parental genotypes are labeled P1 and P2, and the two recombinant genotypes are labeled P3 and P4.

LCIB genotype	*LCI15* genotype	Zeocin resistant (+) / sensitive (–)	Paromomycin resistant (+) / sensitive (–)	Suppressor (+)/ *air dier* (–)	Progeny type
pmp1	WT	–	–	–	P1
ad1	Δ	+	+	+	P2
pmp1	Δ	–	+	+	P3
ad1	WT	+	–	–	P4

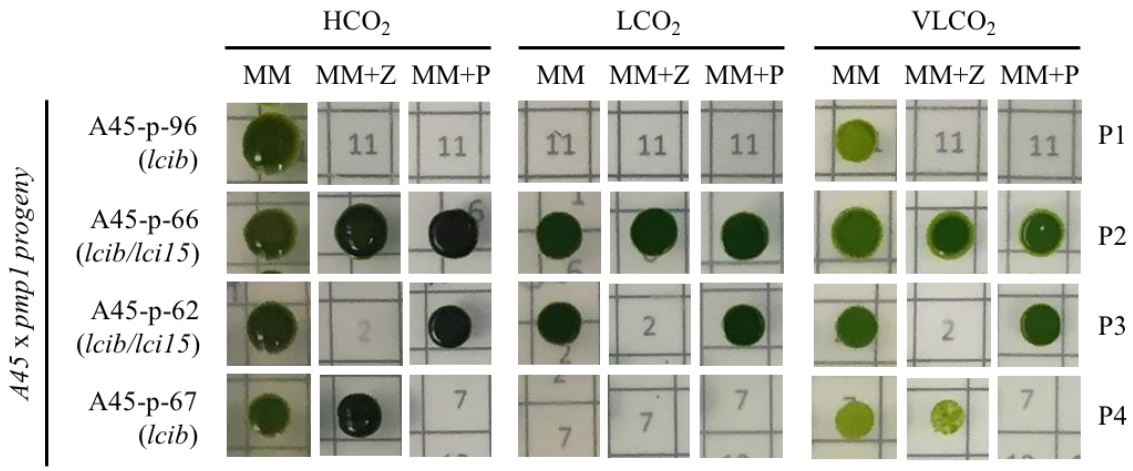


Figure 14. Spot tests of example progeny of the cross between *LCIB-LCI15* double mutant A45 and *pmp1*: one progeny of each of the four possible genotypes (P1–P4) of progeny are shown here.

Of 48 bona fide progeny from crosses of all four *LCIB-LCI15* double mutants with *pmp1*, all zeocin-sensitive (i.e., containing the *pmp1 LCIB* allele) progeny that also

contained an *LCI15* disruption exhibited a suppressor phenotype. Conversely, all progeny lacking an *LCI15* disruption exhibited a non-suppressed, *air dier* phenotype (Figure 14 and Figure 15). These results confirmed cosegregation of the *LCI15* disruption caused by the inserts with the suppressor phenotype.

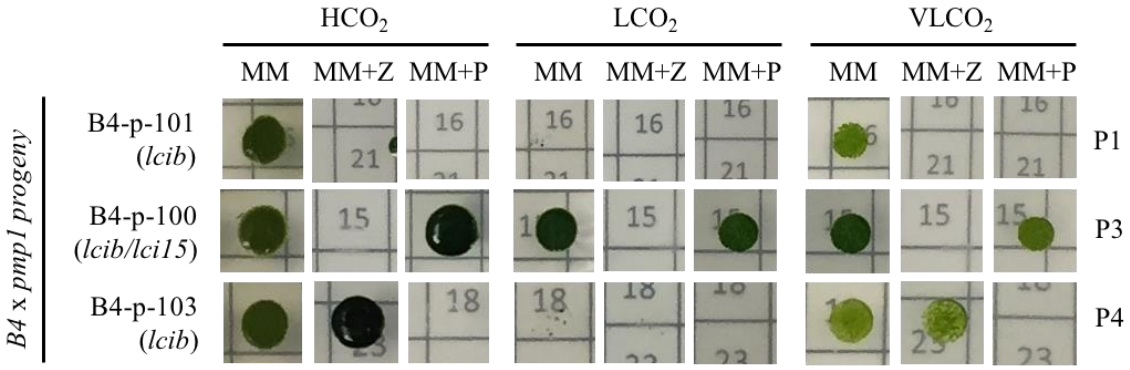


Figure 15. Spot tests of example progeny of the cross between *LCIB-LCI15* double mutant B4 and *pmp1*: one progeny of each of three possible genotypes – P1, P3, and P4 – of progeny are shown here. There were no P2-type progeny from this cross that could be verified as being bona fide progeny.

Knockdown of *LCI15* in *pmp1* using artificial microRNA

If a mutation in *LCI15* is responsible for the suppressor phenotype of the *su1* lines, then we reasoned that *su1* knockdown mutants in a *pmp1* background should generate suppression or at least partial suppression of the *air dier* phenotype.

Of about 300 transformants, 4 Par^R (paromomycin resistant) transformants showed full suppression and 5 showed intermediate suppression of the *pmp1 air dier* phenotype in LCO_2 (Figure 17). However, over time during detailed analysis, most of the transformants lost their resistance to paromomycin and/or their ability to grow in LCO_2. One transformant, A#6, that retained its suppressor phenotype was found to have a single

base change (A > T) at *LCI15* nucleotide position +1854 (downstream of the start codon), which is the penultimate base of intron 4 in *LCI15*. A comparison of cDNA sequences obtained by RT-PCR for the WT and A#6 transcript sequences revealed a 9-bp deletion at the 5' end of *LCI15* exon 5 in A#6 (nucleotides 1–9 of *LCI15* exon 5; +2085 to +2093 downstream of the start codon) as shown in Figure 16.

Figure 16. Sequence comparison of partial *LCI15* coding sequences (CDS) from wild-type and A#6 strains, as determined by RT-PCR. The numbers to the left of the sequences are the nucleotide positions (WT *LCI15* CDS has 951 nucleotides). Partial sequences of exons 4 and 5 are highlighted with red and blue lines, respectively, above the sequence. The blue asterisks under the 5' end of the exon 5 sequence highlight the 9-bp deletion in the A#6 sequence.

This deletion in the transcript sequence is likely due to disruption of the 3' GU–AG consensus sequence (which is required for proper intron splicing) by deletion of the base A in this consensus sequence. With the first 9 bases of exon 5 being 5'-GCGTTGAAG-3', it seems plausible that the lariat structure that forms in the spliceosome recognizes bases 8 and 9 of exon 5 (the AG underlined in the sequence stated here) as the 3' acceptor site, thus splicing out all those first 9 bases as if they were in the 3' end of the intron. This 9-base deletion in the transcript of A#6 results in a

deletion of the first 3 amino acid codons of exon 5 as well, thus altering the predicted amino acid sequence of the LCI15 protein (see Figure 11).

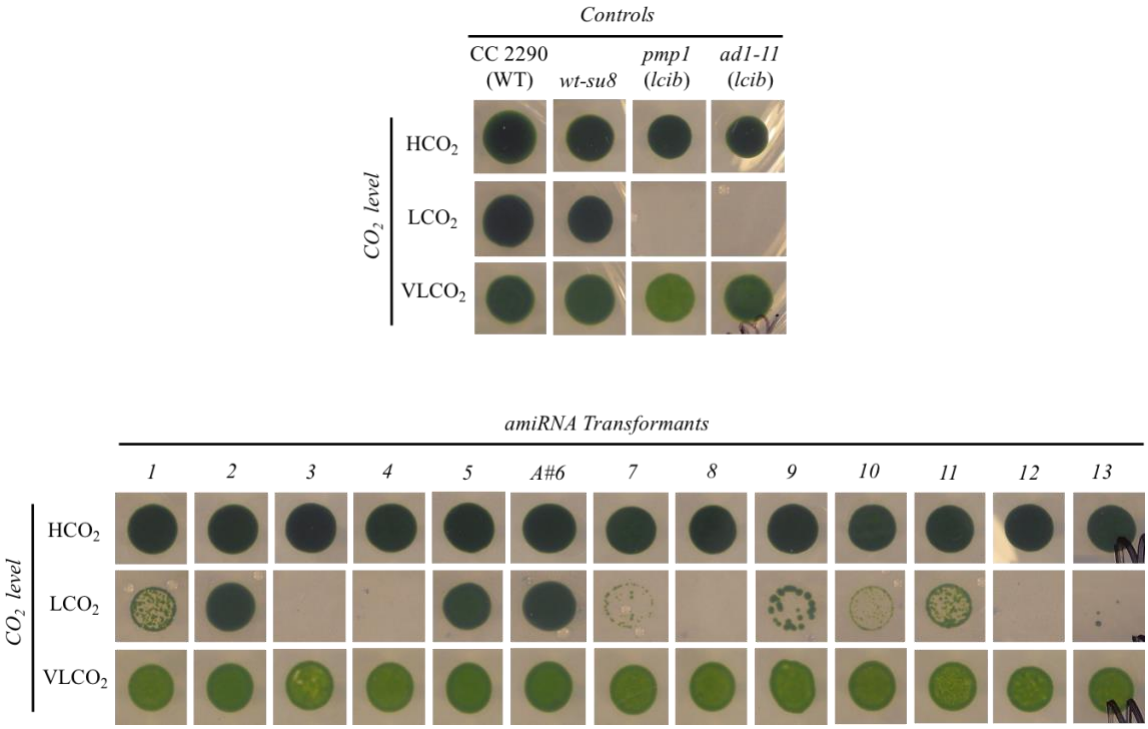


Figure 17. Spot tests of *Par*^{R} (paromomycin-resistant) *pmp1* transformants with *LCI15*-targeting amiRNA and relevant control strains.

Although no bona fide knockdowns of *LCI15* could be confirmed to test the prediction that such knockdowns should result in suppression of the *air dier* phenotype, these experiments did generate another *LCI15* mutant with an *air dier* phenotype, supporting the hypothesis that *LCI15* mutations are responsible for the *su1* suppressor phenotype.

Complementation of *pmp-su1* with WT *LCI15*

In a direct approach to verify whether *su1* is an *LCI15* allele, WT cDNA and genomic DNA under the control of different constitutive promoter/terminator constructs were transformed into *pmp-su1* to try to complement the putative mutant *su1* allele (Table 4).

Table 4. The different vectors and *LCI15* gene inserts used toward complementation of the *LCI15* mutation in *LCIB-LCI15* double mutants.

Vector name	Promoter/ Terminator	*LCI15* Insert Sequence Used for Complementation	Reference (for the Original Vector)
pChlamiRNA3	PSAD	cDNA	Molnar *et al.* (2009)
pChlamy_4	HSP70-RbcS2	cDNA	Rasala *et al.* (2012)
pChlamy_4	HSP70-RbcS2	gDNA	Rasala *et al.* (2012)
pDW2855	HSP70-RbcS2	gDNA	NA
pChlamy_4	HSP70-RbcS2	gDNA + 3' V5 epitope tag	Rasala *et al.* (2012)
pGenD_AphVIII	PSAD	cDNA	Fischer and Rochaix (2001)
pGenD_AphVIII	PSAD	cDNA + N-terminal Strep tag	Fischer and Rochaix (2001)
pGenD_AphVIII	PSAD	cDNA + C-terminal Strep tag	Fischer and Rochaix (2001)
pGenD_AphVIII	PSAD	cDNA + 2x N-terminal FLAG tag	Fischer and Rochaix (2001)

However, we were unable to recover transformants that both reverted to the *air dier* phenotype *and* from which an intact *LCI15* insert (including promoter–coding-sequence–termination sequence) in the genome could be amplified. Some transformants with an *air dier* phenotype were recovered but the *LCI15* transgene was not detectable either by PCR or by Southern blots, indicating that the transgene was not integrated completely in the genome and that the *air dier* phenotype may have resulted from disruption of one or more CCM proteins (or proteins essential for survival in LCO_2).

Other transformants were recovered from which the transgene insert could be demonstrated through PCR, but they did not exhibit an *air dier* phenotype. This suggests either that *SU1* is not an *LCI15* allele, or that there was insufficient expression of functional LCI15 to provide complementation in these transformants, even though the exogenous *LCI15* copy was integrated into the chromosome.

Discussion

Forward genetic screening is a popular approach to isolate mutants that exhibit interesting phenotypes, such as those involving growth, gene regulation, or even survival of an organism. Identifying the causal mutation(s) after collecting mutants, however, can be a tedious task. Two techniques often used in such screens are insertional mutagenesis and ethyl methanesulfonate (EMS) treatment. If the former technique results in an insertion in the genome that cosegregates with the mutant phenotype, identifying the relevant gene locus can, in some cases, be straightforward, given a sequenced genome and methods such as inverse PCR (Ochman *et al.*, 1988) and RESDA-PCR (González-Ballester *et al.*, 2005). However, mutations that leave relatively inconspicuous

"footprints" in the genome, such as those induced by EMS (which are typically random point-mutations) or UV mutagenesis, are harder to pinpoint.

The emergence of next-generation sequencing technologies has streamlined sequencing of entire genomes and the identification of marker systems based on sequence variations across strains or species (Schneeberger, 2014). Single nucleotide polymorphism (SNP) genotyping – identifying SNPs in genome sequences between members of a species – is one such system. Lister *et al.* (2009) describe a strategy that takes advantage of SNP variations between a wild-type strain and a strain with a homozygous recessive mutation in *Arabidopsis thaliana*: DNA from each plant in an F2 population with the mutant phenotype is deep sequenced, and the region of the sequenced data where the difference between the mutant and wild-type SNP densities is equal to 1 (or almost equal to 1) is identified; this region is where the mutated gene is located.

In this project, we applied the same principle to identify the *SU1* locus, albeit with a modification: instead of isolating progeny that expressed the mutant phenotype, we isolated progeny that segregated *against* it. This was because (i) the *su1* mutation renders a wild-type LCO_2 growth phenotype, whether combined with a wild-type or an *air dier LCIB* allele (here, the *pmp1* allele; see Table 1) – mutant progeny would have needed additional PCR and sequencing work to ensure they were in fact progeny and not a vegetative colony mistaken for progeny; and (ii) the growth phenotype of the S1 D2 parental genotype progeny (wild-type in both the *LCIB* and the *SU1* loci) and of progeny carrying the *su1* mutation are indistinguishable using spot tests and would have also needed further sequencing analyses to ensure the presence of the *su1* allele. Screening for *air dier* offspring alone ensured both that they were authentic recombinant progeny *and*

that the allele at the *SU1* locus was known. The *LCIB* allele was *pmp1*, and, although not absolutely necessary, uniformity in all the offsprings' *LCIB* allele was helpful since it served as a control for SNP density ratios in a region other than the one that was of interest to us.

S1 D2 SNP densities, which were measured as a ratio of the occurrence of non-WT SNPs in the total number of sequencing reads at each nucleotide position in the pooled DNA, were calculated for every position where there was a SNP call (i.e., a SNP difference between the reference strains and S1 D2 was detected). The averages of all calculated densities within 10,000-bp regions were plotted against the 10,000-bp windows of the entire genome (Figure 10). With the peak of the graph still covering a fairly broad region of the genome (1.5M bps encompassing about 245 genes), SNP differences that one would expect in the *SU1* locus between *pmp-su1* and the reference strains and between *pmp-su1* and S1 D2 as well as those that resulted in changes to predicted amino acid sequences were used to reduce the number of possible candidate genes with this region. One of two genes identified, *LCI15*, resulted in a change in the predicted amino acid change in *pmp-su1*, and was an attractive candidate gene since, as a low CO_2-inducible gene, it was apparently responsive to LCO_2, and was regulated by CIA5/CCM1, the master regulator of the CCM.

Subsequent to identification of *LCI15* by deep sequencing the pooled DNA and then finding putative genes based on S1 D2-to-reference strains SNP ratios, complementation of the *su1* mutant with a wild-type *LCI15* genomic or cDNA insert (under the control of different constitutive promoter-terminator sequences) did not prove fruitful, likely because the selection was for a negative trait, i.e., successful expression of

the transgene would have resulted in cell lines that were *air dier*-suppressors to revert to the *air dier* phenotype, since diploid analyses showed that the *su1* mutation is recessive in nature (Duanmu and Spalding, 2011). This negative selection, in combination with the notorious difficulty of expressing an introduced gene in Chlamydomonas, even if the introduced gene is from Chlamydomonas itself (León and Fernández, 2007; Schroda *et al.*, 2000), likely rendered screening for complementation of the *LCI15* mutation nearly impossible, or at the very least, highly improbable. In some transformants, the entire *LCI15* transgene was detectable by PCR-amplification, but there was still no detectable complementation (i.e., they retained wild-type growth in LCO_2); this may be explained by complete or partial gene-silencing resulting in inadequate levels of the transgenic LCI15 protein.

Previous studies have explored possible reasons for poor expression levels of nuclear transgenes and even the complete loss of their expression in the long-term; some reasons being epigenetic silencing mechanisms such as histone modifications; RNA interference processes; and transgene integration position in the genome (Cerutti *et al.*, 1997; Kim *et al.*, 2015; Rosales-Mendoza *et al.*, 2012; Schroda, 2006; van Dijk *et al.*, 2007). To address possible causes for silencing of wild-type *LCI15* gene expression, studies of the encoded protein's function would be required (please see Chapter 3). Assuming that the absence of the LCI15 protein results in the activation of an LCIB-like protein in an *LCIB* mutant background, or in the activation of an alternate pathway in order to enable suppression, it is plausible that the wild-type gene confers at least a small negative selection even when grown in TAP.

While the complementation experiments using different vectors and *LCI15* inserts (listed in Table 4) did not result in a reversion to the *air dier* phenotype in *pmp-su1*, and, therefore, did not appear to successfully complement the *su1* mutation, we could have concluded that a mutation in *LCI15* was not the one responsible for the *air dier*-suppressor phenotype in *pmp-su1*.

Table 5. List of *LCIB-LCI15* mutants that exhibit an *air dier*-suppressor phenotype. Strains with an asterisk (*) next to their name were used in cosegregation experiments to confirm that *su1* was a mutation in *LCI15*. Par^R: paromomycin resistance. Zeo^R: zeocin resistance.

Strain name	Background strain	*LCI15* mutation	Antibiotic resistance	Reference
pmp-su1	*pmp1*	Exon 4, base 2: G > T base change	--	Duanmu and Spalding (2011)
A#6	*pmp1*	Intron 4, base 198: A > T base change	Par^R	This dissertation
Cas9-27*	*pmp1*	Exon 2, bases 240–246: 7bp deletion	--	This dissertation
Cas9-28	*pmp1*	Exon 2, base 266: 1bp deletion	--	This dissertation
Cas9-19	*pmp1*	Exon 2, bases 205–206: 2bp deletion	--	This dissertation
A45*	*ad1*	*AphVIII* insert in 5’ half of the gene	Zeo^R, Par^R	This dissertation
B4*	*ad1*	Partial *AphVIII* insert in 3’ half of the gene	Zeo^R, Par^R	This dissertation
B24*	*ad1*	Insert in 3’ half of the gene	Zeo^R, Par^R	This dissertation
B28*	*ad1*	Insert in 5’ half of the gene	Zeo^R, Par^R	This dissertation

However, over time, we had amassed eight different mutants (not including the original mutant, *pmp-su1*) in *LCI15* in *LCIB*-mutant backgrounds, all of which exhibited suppression of the lethal phenotype in LCO_2 (Table 5). Although this was not conclusive evidence that the *su1* mutation was in *LCI15*, it seemed highly improbable that it was not.

In the absence of the ability to complement the *LCI15* mutation, the classical approach of cosegregation analysis was employed to verify that the mutation in *su1* was responsible for the suppression phenotype. Five of the 8 additional *LCIB-LCI15* mutants were crossed to either *pmp1* or *ad1* (*LCIB* allelic mutants) and their progeny screened for cosegregation of the *LCI15* mutation and the *air dier*-suppressor phenotype. In total, 81 progeny were confirmed as bona fide progeny from all these crosses, and all of these progeny that carried the *LCI15* mutation were able to survive in LCO_2, and those that inherited the wild-type *LCI15* allele were unable to grow in LCO_2, thus establishing that the *LCI15* mutation is responsible for the suppressor phenotype.

Thus, using the deep sequencing approach and cosegregation analysis by crosses, we were able to successfully identify and confirm that the *su1* mutation was in the *LCI15* gene. In the next chapter, we discuss putative functions of the LCI15 protein and how its absence might be resulting in survival of *LCIB* mutants in LCO_2.

References

Cerutti, H., Johnson, A.M., Gillham, N.W., & Boynton, J.E. 1997. Epigenetic silencing of a foreign gene in nuclear transformants of *Chlamydomonas*. *Plant Cell* 9:925-945.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., McKenna, A., Fennell, T.J., Kernytsky, A.M., Sivachenko, A.Y., Cibulskis, K., Gabriel, S.B., Altshuler, D., & Daly, M.J. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 43:491-498.

Dodds, W.K., & Whiles, M.R. 2010. Chapter 13 - Carbon. Pp. 323-343 in, *Freshwater Ecology (Second Edition)*. Academic Press, London. p 323-343.

Duanmu, D., & Spalding, M.H. 2011. Insertional suppressors of *Chlamydomonas reinhardtii* that restore growth of air-dier *lcib* mutants in low CO_2. *Photosynth Res* 109:123-132.

Duanmu, D., Wang, Y., & Spalding, M.H. 2009. Thylakoid lumen carbonic anhydrase (*CAH3*) mutation suppresses air-dier phenotype of *LCIB* mutant in *Chlamydomonas reinhardtii*. *Plant Physiol* 149:929-937.

Evans, J.R., & Von Caemmerer, S. 1996. Carbon Dioxide Diffusion inside Leaves. *Plant Physiology* 110:339-346.

Fischer, N., & Rochaix, J.D. 2001. The flanking regions of *PsaD* drive efficient gene expression in the nucleus of the green alga *Chlamydomonas reinhardtii*. *Mol Genet Genomics* 265:888-894.

Gallaher, S.D., Fitz-Gibbon, S.T., Glaesener, A.G., Pellegrini, M., & Merchant, S.S. 2015. Chlamydomonas Genome Resource for Laboratory Strains Reveals a Mosaic of Sequence Variation, Identifies True Strain Histories, and Enables Strain-Specific Studies. *Plant Cell* 27:2335-2352.

Geraghty, A.M., Anderson, J.C., & Spalding, M.H. 1990. A 36 Kilodalton Limiting-CO_2 Induced Polypeptide of *Chlamydomonas* Is Distinct from the 37 Kilodalton Periplasmic Carbonic Anhydrase. *Plant Physiol* 93:116-121.

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.

Gorman, D.S., & Levine, R.P. 1965. Cytochrome f and plastocyanin: their sequence in the photosynthetic electron transport chain of Chlamydomonas reinhardi. *Proc Natl Acad Sci U S A* 54:1665-1669.

Gross, C.H., Ranum, L.P., & Lefebvre, P.A. 1988. Extensive restriction fragment length polymorphisms in a new isolate of *Chlamydomonas reinhardtii*. *Curr Genet* 13:503-508.

Harris, E.H. 1989. The *Chlamydomonas* Sourcebook: A Comprehensive Guide to Biology and Laboratory Use. Academic Press, San Diego, CA.

Jiang, W.Z., & Weeks, D.P. 2017. A gene-within-a-gene Cas9/sgRNA hybrid construct enables gene editing and gene replacement strategies in *Chlamydomonas reinhardtii*. *Algal Research*.

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 U S A* 113:14716-14721.

Kathir, P., LaVoie, M., Brazelton, W.J., Haas, N.A., Lefebvre, P.A., & Silflow, C.D. 2003. Molecular map of the *Chlamydomonas reinhardtii* nuclear genome. *Eukaryot Cell* 2:362-379.

Kim, E.J., Ma, X., & Cerutti, H. 2015. Gene silencing in microalgae: mechanisms and biological roles. *Bioresour Technol* 184:23-32.

León, R., & Fernández, E. 2007. Nuclear transformation of eukaryotic microalgae: historical overview, achievements and problems. *Adv Exp Med Biol* 616:1-11.

Li, H., & Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.

Lister, R., Gregory, B.D., & Ecker, J.R. 2009. Next is now: new technologies for sequencing of genomes, transcriptomes, and beyond. *Curr Opin Plant Biol* 12:107-118.

Molnar, A., Bassett, A., Thuenemann, E., Schwach, F., Karkare, S., Ossowski, S., Weigel, D., & Baulcombe, D. 2009. Highly specific gene silencing by artificial microRNAs in the unicellular alga *Chlamydomonas reinhardtii*. *Plant J* 58:165-174.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M.A. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297-1303.

Ochman, H., Gerber, A.S., & Hartl, D.L. 1988. Genetic applications of an inverse polymerase chain reaction. *Genetics* 120:621-623.

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:e43349.

Ricklefs, R.E., & Miller, G.L. 2000. Ecology. W. H. Freeman.

Rosales-Mendoza, S., Paz-Maldonado, L.M., & Soria-Guerra, R.E. 2012. *Chlamydomonas reinhardtii* as a viable platform for the production of recombinant proteins: current status and perspectives. *Plant Cell Rep* 31:479-494.

Schneeberger, K. 2014. Using next-generation sequencing to isolate mutant genes from forward genetic screens. *Nat Rev Genet* 15:662-676.

Schroda, M. 2006. RNA silencing in *Chlamydomonas*: mechanisms and tools. *Curr Genet* 49:69-84.

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.

Shimogawara, K., Fujiwara, S., Grossman, A., & Usuda, H. 1998. High-efficiency transformation of *Chlamydomonas reinhardtii* by electroporation. *Genetics* 148:1821-1828.

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.

Spalding, M.H., Spreitzer, R.J., & Ogren, W.L. 1983. Reduced Inorganic Carbon Transport in a CO_2-Requiring Mutant of *Chlamydomonas reinhardii*. *Plant Physiol* 73:273-276.

Umen, J.G., & Olson, B.J.S.C. 2012. Chapter Six - Genomics of Volvocine Algae. Pp. 185-243 in: Gwenaël, P., (ed), *Advances in Botanical Research*. Academic Press. p 185-243.

van Dijk, K., Xu, H., & Cerutti, H. 2007. Epigenetic silencing of transposons in the green alga *Chlamydomonas reinhardtii*. Pp. 159 – 178 in: Nellen, W., & Hammann, C., (eds), *Small RNAs:: Analysis and Regulatory Functions*. Springer Berlin Heidelberg, New York. p 159 – 178.

Vance, P., & Spalding, M.H. 2005. Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Canadian Journal of Botany* 83:796-809.

Vysotskaia, V.S., Curtis, D.E., Voinov, A.V., Kathir, P., Silflow, C.D., & Lefebvre, P.A. 2001. Development and characterization of genome-wide single nucleotide polymorphism markers in the green alga *Chlamydomonas reinhardtii*. *Plant Physiol* 127:386-389.

Wang, Y., & Spalding, M.H. 2006. An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 103:10110-10115.

Wang, Y., & Spalding, M.H. 2014a. 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.

Wang, Y., & Spalding, M.H. 2014b. LCIB in the *Chlamydomonas* CO_2-concentrating mechanism. *Photosynth Res* 121:185-192.

Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S., 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.

CHAPTER 3. FUNCTIONAL ANALYSIS OF LCI15

Introduction

Photosynthesis, a process that occurs in most plants, algae, and cyanobacteria, contributes to the oxygen (O_2) levels in the Earth's atmosphere and thus supports sustenance of aerobic life on the planet. This process uses carbon dioxide (CO_2) in the light-independent reactions (also known as the Calvin-Benson-Bassham (CBB) cycle or the Calvin cycle) and produces O_2 through the light-dependent reactions. The first step of the Calvin cycle is carried out by arguably one of the most important enzymes in the world, ribulose-1,5-bisphosphate carboxylase/oxygenase (commonly referred to as Rubisco), which combines CO_2 and a five-carbon compound, ribulose-1,5-bisphosphate (RuBP), to form a six-carbon compound; this cycle eventually leads to the production of glucose. However, Rubisco is capable of utilizing both CO_2 and O_2 as substrates: when CO_2 is fixed, the Calvin cycle commences and there is a net gain of six carbon atoms, but when O_2 is fixed, the wasteful photorespiratory process is pursued and there is a net loss of three fixed carbon atoms (Buchanan *et al.*, 2015).

In order to overcome the competition posed by O_2 and to promote the Rubisco RuBP-carboxylation reaction, a number of photosynthetic organisms have developed CO_2-concentrating mechanisms (CCMs); in many terrestrial plants these are known as the C4 and CAM (Crassulacean Acid Metabolism) systems (Dodd *et al.*, 2002; Furbank *et al.*, 2000; Sage, 2004). The goal of these mechanisms is to increase concentration of CO_2 around Rubisco, thus minimizing photorespiration. These mechanisms also help overcome the limitation of the slow catalytic rate of Rubisco.

Inorganic carbon (Ci) in aquatic environments is present mainly in the form of CO_2 or HCO_3^- (their relative proportion depends mainly on the ambient pH and the solubility of CO_2), and carbonic anhydrases (CAs) are needed for interconversion between the Ci species, largely because Rubisco can use only CO_2 (and not HCO_3^-) as a substrate. Aquatic environments pose an additional deterrent to carbon fixation in photosynthesis: the diffusion of CO_2 in water is 10,000 times slower than it is in air (Winck *et al.*, 2013).

Chlamydomonas reinhardtii (henceforth referred to as "Chlamydomonas"), a unicellular aquatic green alga, has a CCM that is induced in ambient CO_2 levels known as 'limiting CO_2', which includes low CO_2 (LCO_2 or air-level CO_2; 0.03–0.4% CO_2) and very low CO_2 ($VLCO_2$; <0.02% CO_2) (Vance and Spalding, 2005). The system is built on three main pillars that help mitigate the effects of Rubisco's inefficiency and those of the additional disadvantages of aquatic habitats: active inorganic carbon (Ci) uptake systems, carbonic anhydrases (CAs), and the formation of pyrenoids, which are microcompartments within the cell where Rubisco is concentrated. As its name suggests, the CCM helps concentrate CO_2 within the cell, specifically near Rubisco, to facilitate and encourage the carboxylation of RuBP by Rubisco. In order to achieve this, the CCM triggers changes in the physiology and gene expression in the cell.

The entry of Ci into the cell takes place either by diffusion of CO_2 or by active uptake of CO_2 and HCO_3^-. In order to reach Rubisco, which is housed within the chloroplast, Ci has to cross the plasma membrane and the chloroplast envelope. The CAs perform the important function of interconversion between the Ci species in the CCM;

for instance, CAH3, a thylakoid lumen CA, is largely responsible for the dehydration of HCO_3^- to CO_2 close to the pyrenoid, thus providing a CO_2 stream for Rubisco.

LCIB mutants, which lack a plastid protein that is essential for survival only in LCO_2, are known as *air diers*, and two of these are *pmp1* (Spalding *et al.*, 1983) and *ad1* (Wang and Spalding, 2006). For a better understanding of the role of LCIB in the CCM, Duanmu and Spalding (2011) generated multiple *air dier*-suppressor lines with second-site mutations, and the location of one of these, *SU1* in the *pmp-su1* line, was determined by deep sequencing, as described and discussed in Chapter 2.

Now that we have identified the locus corresponding to *SU1* and have demonstrated that mutations in *LCI15* are responsible for suppression of the *air dier* phenotype of *LCIB* mutants, it is important to determine how *LCI15* mutations are able to suppress this phenotype. We first need to determine *how* Chlamydomonas may be able to circumvent or bypass the apparent absolute need for LCIB function in air-level CO_2; specifically, how Ci uptake in Chlamydomonas has changed in *LCI15* mutants to enable photosynthesis and growth in air-level CO_2. We also need to address the much more difficult question of how mutations in *LCI15* cause the changes in Ci uptake, including asking the obvious question of whether the nature of LCI15 offers insight into how it might function and, therefore, how the absence of its function may significantly affect Ci uptake.

With regard to changes in the pathway or mechanism of Ci uptake that suppress the conditional-lethal, *air dier* phenotype of *LCIB* mutants, some of the possible scenarios by which LCIB might be replaced or bypassed in LCO_2 include (i) the formation of a LCIB/LCIC-like complex that replaces the original in function;

(ii) upregulation of an alternate pathway that substitutes for or bypasses the need for LCIB in the CCM; (iii) relief of the CO_2 inhibition of the HLA3-LCIA-associated HCO_3^- pathway observed in LCO_2; and (iv) increase in the production of a component(s), not CCM-related and required for growth, such as a general increase in photosynthetic efficiency that obviates the need for the increased concentration of CO_2 in air-level CO_2.

Based on the fact that *air dier*, *LCIB* mutants have a lethal growth phenotype in LCO_2, it has been determined that LCIB is an important component of the CCM and is required for survival in this ambient CO_2 growth level. As expected, it is vital for active CO_2 uptake in LCO_2, and recent physiological studies (Wang and Spalding, 2014a) found that it plays a role in Ci uptake in $VLCO_2$ as well. Apart from its involvement in active CO_2 uptake, LCIB is also believed to prevent escape of CO_2 (that has not been fixed by Rubisco) from the pyrenoid inside the cell. A current line of thought, supported by a study of the LCIB structure (Jin *et al.*, 2016), is that it prevents this escape by converting the CO_2 into HCO_3^-, which cannot diffuse out of the cell as easily as CO_2 can.

There are multiple LCIB-like genes that comprise the LCIB gene family in Chlamydomonas – *LCIB*, *LCIC*, *LCID*, and *LCIE* – and all of these genes are constitutively expressed in HCO_2 and upregulated by limiting CO_2 (Fang *et al.*, 2012; Miura *et al.*, 2004; Wang and Spalding, 2006; Yamano *et al.*, 2008). *LCIB* and *LCIC* accumulated similar levels of transcript, and, in comparison, those of *LCID* and *LCIE* are moderate and low, respectively. LCIB and LCIC form a 350kDa hexameric complex that localizes to the periphery of the pyrenoid (also called the peripyrenoid area) in $VLCO_2$ and is distributed throughout the chloroplast stroma in both HCO_2 and LCO_2 (Wang and Spalding, 2014b).

One hypothesis for how the effect of the *LCIB* mutation may be masked in a *LCIB*-*LCI15* mutant is by the formation of an LCIB/LCIC-like complex which would replace the original's function, based on the premise that the absence of LCI15 is directly or indirectly allowing for the formation of such a replacement complex. The other proteins in the LCIB family would be prime candidates for this, functioning alone as monomers or homomultimers, or as heteromultimers with another member of the LCIB family.

Previous work has shown that LCO_2-acclimated *LCIB* mutant cells express *LCIC* at the mRNA level, but not at the protein level (Yamano *et al.*, 2010). If *LCI15* inhibits expression of one or more of *LCIC*, *LCID*, and *LCIE*, *LCIB*-*LCI15* mutants might exhibit an increase in the transcript (and, perhaps, protein) levels of these genes. If the encoded proteins can accumulate and possibly form heteromultimers with one another (for example, an LCIC/LCID or LCID/LCIE complex), they might successfully take over the roles of LCIB.

LCIA, a chloroplast envelope Chlamydomonas nitrate-transporter-related protein, has been reported to be involved in active HCO_3^- transport along with HLA3, a plasma membrane protein from the ATP-binding cassette transporter family (Duanmu *et al.*, 2009; Wang and Spalding, 2014a). Photosynthetic oxygen evolution analyses of *LCIB* and *LCIA*-*LCIB* mutants by Wang and Spalding (2014a) revealed that this LCIA-associated system is inhibited in air-level CO_2, which explained why this system does not normally compensate for the LCIB deficiency in *LCIB* mutants in LCO_2 levels of growth. It was surmised that CO_2, either directly or indirectly through other proteins, is responsible for the inhibition. Since Ci uptake in LCO_2 is largely accounted for by the

LCIB-associated CO_2 pathway, and not by the LCIA-associated HCO_3^- pathway, it is plausible that removal of the CO_2-inhibition of the LCIA/HCO_3^- system might restore growth (partial or otherwise) of *LCIB* mutants in air-level CO_2. If LCI15 plays a regulatory role and affects the inhibition of the LCIA-related Ci uptake system, a mutation in *LCI15* might result in withdrawal of that inhibition, establishment of an alternative to the LCIB-associated Ci uptake pathway, and suppression of the *air dier* phenotype.

A third scenario might be that an alternate CO_2 uptake pathway is upregulated in the absence of LCI15. This could involve LCI1, a putative Ci transporter localized to the plasma membrane, and/or other CCM components, such as the LCIA/HCO_3^- pathway or unknown proteins comprising an as yet unidentified Ci uptake pathway.

To determine if one of these possibilities could explain how an *LCI15* mutation offsets the detrimental effect that an *LCIB* mutation has on growth in LCO_2, relative transcript and protein levels of various genes known to be involved in the CCM were examined in a wild-type (WT) strain, and *LCIB*, *LCI15*, and *LCIB-LCI15* mutants. We found that a number of CCM genes have higher transcript abundance in *LCI15* mutants, and that LCI15 likely plays an overarching role in the regulation of these genes.

This also raises the question of *how* LCI15 effects this change: the absence of the protein triggers an increase in the expression of these genes, or, inversely, its presence prompts a decrease in their expression. Considering that the difference between the *LCI15* mutant and non-mutant lines is seen at the transcript level (and subsequently also at the protein level for those proteins that we could probe), LCI15 might be involved in regulation at the transcriptional or post-transcriptional level. If not a negative regulator

(such as a repressor or a corepressor) itself, it could be a component that is important for the proper functioning of a negative regulatory complex, or a part of a gene silencing process (such as RNA interference) that limits expression of some genes by targeting and neutralizing their mRNA.

Materials and Methods

Cells Strains and Growth Conditions

The strains used in this report were CC-3269 (also known as 2137; Spreitzer and Mets (1981)) as a wild-type control; *pmp1*-7k+, an *LCIB* mutant progeny from eleven backcrosses between 2137 and CC-4676 (also known as 16-5K, a *pmp1 air dier* mutant; Spalding *et al.* (1983)); Cas9-27, Cas9-28, and Cas9-19, three *LCIB-LCI15* double mutants in a *pmp1* background; *wt-su1*, an *LCI15* mutant progeny generated from a cross between Cas9-27 and CC-621 (a high mating efficiency WT strain) and has a wild-type *LCIB* allele; and A#6, an *LCIB-LCI15* mutant in the *pmp1* background with a three-residue deletion in its predicted amino acid sequence. HCO_2-grown cultures were grown to about 500,000 cells/ml in liquid minimal medium, then bubbled with HCO_2 (5% [v/v] CO_2 enriched with air) until mid-log phase of growth and then the cells were harvested. LCO_2-induced cultures were first grown to about 1 million cells/ml in HCO_2, then bubbled with LCO_2 (~0.04% CO_2) for 15–18 hours in order to induce the CCM, following which they were harvested.

Identification of LCI15 Homologs and Predicted Domains

BLAST (Basic Local Alignment Search Tool) searches were performed on the NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi) and JGI Phytozome

(https://phytozome.jgi.doe.gov/pz/portal.html) websites to identify proteins homologous to LCI15 and to PRL1 in Chlamydomonas. Protein domain and subcellular localization prediction searches were performed using InterPro (https://www.ebi.ac.uk/interpro/) and PSORTb (http://www.psort.org/psortb/index.html).

Protein Sequence Alignments

Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) was used to generate multiple sequence alignments. ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/) was then used to visualize the alignment and similarities between residues (Robert and Gouet, 2014).

RNA Isolation and Reverse-Transcriptase PCR

Reverse-transcriptase PCR (RT-PCR) was performed on RNA isolated from HCO_2- and LCO_2-acclimated strains to amplify various CCM gene transcripts. Cultures (about 50 ml of mid-log phase concentration) were centrifuged at 10,000 *g*, the supernatant (growth medium) was removed, and the cells were snap frozen in liquid nitrogen before being stored at –80°C until the next step. RNA was isolated using the RNeasy® Plant Mini Kit (Qiagen, Inc.) according to the manufacturer's protocol, with the exception that the initial cell lysis step was performed by vortexing the cells for about 1–2 minutes. The RNA obtained was then cleaned of any contaminating DNA using the Ambion® TURBO DNA-free™ kit (Life Technologies Corporation). Libraries of cDNA for each strain were generated using the Invitrogen SuperScript® III First-Strand Synthesis System (Life Technologies Corporation) using about 600 ng of RNA, measured

using a NanoDrop® ND-2000 Spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE). The cDNA was diluted to an equivalent of that which would be derived from 1 ng/µl RNA, and 3 µl of the dilution was used as template for PCR using GoTaq® Green Master Mix (Promega Corporation). The RT-PCR was performed at four PCR cycle numbers: 25, 30, 35, and 40 cycles. The PCR products were then electrophoresed on 1.5–2% agarose gels. CBLP (*Chlamydomonas* β-subunit-like protein) and UBC8 (a ubiquitin ligase; Phytozome ID: Cre03.g159200; Jokel *et al.*, 2015) were used as reference genes.

Protein Extraction and Western Immunoblotting

Western immunoblots were performed on protein extracts from HCO_2- and LCO_2-acclimated strains to detect various CCM proteins. Harvested cultures (about 7 ml of mid-log phase grown cells) were centrifuged to remove the growth medium, and the precipitated cells were resuspended to visually similar concentrations using SDS lysis buffer (10mM Tris-HCl [pH 7.5], 1mM EDTA, 10mM NaCl, 2% SDS) with a dissolved Roche cOmplete™ Protease Inhibitor Cocktail tablet (Mini, EDTA-free; Sigma-Aldrich, Inc.; 1 tablet in 10ml of SDS lysis buffer). The resuspension was then put through three freeze-thaw cycles using liquid nitrogen and warm water. 100 µl from each sample was used for performing the BCA (bicinchoninic acid) assay to measure total protein concentrations with the Pierce BCA Protein Assay kit (Thermo Fisher Scientific, Inc.; Catalog # 23225) following the manufacturer's protocol. The SDS lysis buffer was used as the diluent in the BSA (bovine serum albumin) standards in the BCA assay. The rest of the cell suspension in the SDS lysis buffer was diluted in 4x SDS sample loading buffer

(1M Tris-HCl [pH 6.8], 8% (w/v) SDS, 40% (v/v) glycerol, 14.7M β-mercaptoethanol, 0.5M EDTA, 0.8% (w/v) bromophenol blue; see Sambrook (2001)), heated to 90–100ºC for 5 minutes, and stored in –20º C.

Total protein extracts of 1 or 2 µg (250 ng for detection of large subunit of Rubisco) as measured by the BCA assay were separated on 12% (v/v) SDS-polyacrylamide gels. Specific antibodies were used to detect the various CCM proteins and the two loading control proteins, COX2B (mitochondrial cytochrome c oxidase subunit 2B) and the large subunit of Rubisco (RbcL), using horseradish peroxidase (HRP)-conjugated secondary antibodies and final detection by chemiluminescence with the SuperSignal West Pico system (Thermo Scientific Inc.) The Rubisco antibodies were generous gifts from Dr. Sabeeha Merchant (University of California, Los Angeles), Dr. James Moroney (Louisiana State University), and Dr. Howard Griffiths (University of Cambridge, UK). The COX2B antibody was also a gift from Dr. Merchant.

Results

The Nature of LCI15

LCI15 is composed of 316 amino acids with a predicted molecular weight of 34.34 kDa. It has no obvious, predicted transmembrane domains or organelle-targeting peptides, but recent work by Mackinder *et al.* (2017) showed that LCI15 localizes to the cytosol and the flagella, although since these observations were in $VLCO_2$-acclimated cells, we cannot be sure where the localization might be in LCO_2 or whether there is even a change in localization between LCO_2 and $VLCO_2$. The protein is predicted by InterPro to have a cobalamin biosynthesis protein CobW-like C-terminal domain and a WW (two

tryptophan residues) domain near its C-terminal end (Figure 18), but no other identifiable domains were predicted.

Figure 18. The *LCI15* gene has five coding exons (alternate exons are represented by dark and light green bars (/). The LCI15 protein has two recognizable domains: a cobalamin synthesis protein CobW-like C-terminal domain ( ), and a WW (two tryptophan residues) domain ( ).

CobW is one of about 30 enzymes involved in the biosynthesis of cobalamin (commonly referred to as vitamin B_{12}), a cobalt-containing coenzyme. Bacteria, archaea, animals, and protists have cobalamin-dependent enzymes, but only bacteria and archaea synthesize cobalamin themselves. On the other hand, fungi and plants neither produce nor need it for their metabolic activity (Bertrand *et al.*, 2011; Fang *et al.*, 2017; Martens *et al.*, 2002). CobW proteins contain a histidine-rich C-terminal region, which indicates a possible role in metal binding in chelation systems (Rodionov *et al.*, 2003), and CobW-like proteins have been implicated in activation of nitrile hydratase (Hashimoto *et al.*, 1994). However, it is not clear what the function of CobW domains might be in many proteins where the domain is predicted.

WW domains consist of about 40 amino acids with two conserved tryptophan residues roughly 20-22 amino acids apart. They are found in functionally and structurally diverse and unrelated proteins and are predicted to mediate protein-protein interactions. Examples of the diversity of WW domain-containing proteins include: a putative transcriptional regulator that is highly expressed in human gastric cancer cells (Li *et al.*, 2017); an oxidoreductase that functions as a tumor suppressor in humans (Bunai *et al.*,

2017); and a RNA-binding protein that might act as a posttranscriptional repressor of circadian genes in Chlamydomonas (Matsuo and Ishiura, 2010). WW domains reportedly bind proline-rich sequences (Kay *et al.*, 2000; Smerdon and Yaffe, 2010; Sudol and Hunter, 2000) that contain PPxY and LPxY (Bruce *et al.*, 2008); PPLP (Huang *et al.*, 2009); PPR (Bedford *et al.*, 2000; where P represents a proline residue; Y is tyrosine; L is leucine; and R is arginine); and p-SP (phosphoserine-proline) and p-TP (phosphoserine-threonine) motifs (Lu *et al.*, 1999). So, the presence of a predicted WW domain in LCI15 suggests it is likely to interact with another protein, which is not very informative by itself.

As illustrated in Figure 19, a sequence comparison of the predicted wild-type LCI15 protein and that of *pmp-su1* reveals that the single base change in the first codon of *LCI15* exon 5 results in what would seem to be a fairly benign change at the amino acid level: a glycine to valine substitution in the CobW-like C-terminal domain. With both these amino acids being hydrophobic, it is unclear how this change might result in a non-functional protein or even a protein with significantly altered function. Since no specific targeting sequences were predicted from the protein sequence, it seems unlikely to have resulted in mistargeting of the protein away from the right organelle. It is possible that this amino acid change might lead to misfolding of the final protein, since valine, which has a molecular weight of 99 Da, has a larger side-chain than glycine (57 Da). It is conceivable that the larger size of valine, in combination with its relative location, could negatively affect the three-dimensional conformation of the overall protein or of the CobW-like domain specifically, especially if the CobW-like domain is important for LCI15 function.

The mutant generated from amiRNA targeting of *LCI15*, A#6, has a single nucleotide substitution that led to a 9-bp deletion at the cDNA level and a subsequent 3 amino acid predicted deletion (Figure 19; see Chapter 2 for more information). This deletion (residue positions 269-271) also falls within the CobW-like C-terminal domain and could disturb the ability of the LCI15 protein to function, assuming this domain has some functional significance.

While *pmp-su1* and A#6 have relatively small changes in the nucleotide sequence and the predicted translation product, the three CRISPR/Cas9-generated mutants, Cas9-27, Cas9-28, and Cas9-19 have deletions in their *LCI15* genomic DNA sequences that lead to frame-shift mutations and, subsequently, premature stop codons. Cas9-27 has a 7-bp deletion from +522, Cas9-28 has a 1-bp deletion at +548, and Cas9-19 has a 2-bp deletion at +487-488 (all base positions with respect to the start codon). These deletions result in changes in the predicted amino acid sequences starting at residues 104, 112, and 92, respectively, and the overall protein lengths are only 142, 144, and 157 residues, while wild-type LCI15 has 316 residues (Figure 19). Cas9-19 also has some amino acid insertions in its sequence. A significant difference that stands out is also the complete lack of the CobW-like C-terminal domain and the WW domain in these three mutants.

PRL1 and PRL1-Interacting Factor L

LCI15 is annotated as a homolog to the PRL1-interacting factor L protein in *Arabidopsis thaliana* (hereafter Arabidopsis). While little is known about most PRL1-interacting factors, PRL1 (Pleiotropic regulatory locus 1) itself is a nuclear protein involved in pleiotropically controlling sugar and hormone responses in Arabidopsis.

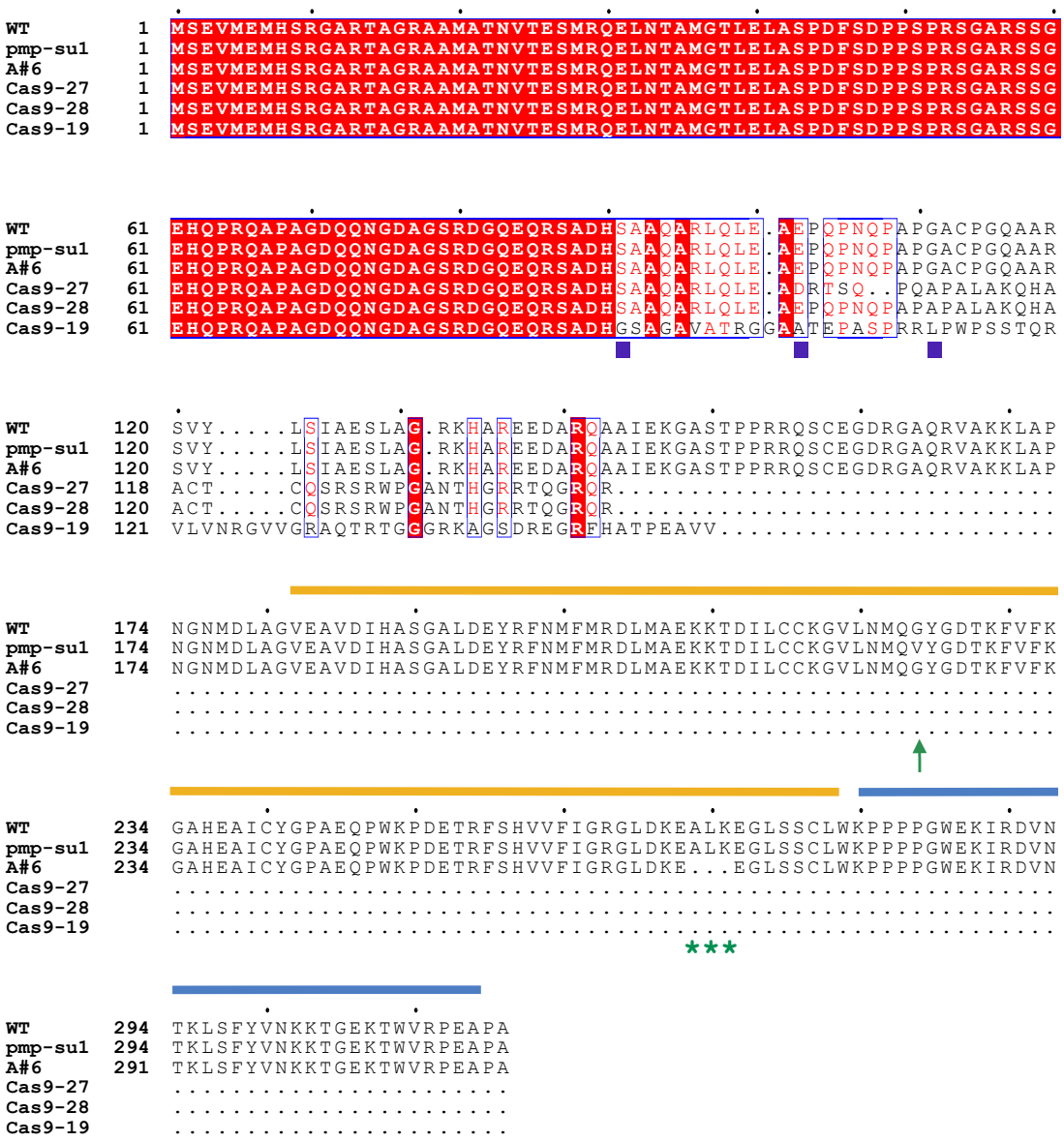


Figure 19. Sequence comparison of predicted LCI15 proteins from wild-type, *pmp-su1*, A#6, Cas9-27, Cas9-28, and Cas9-29 strains. The green arrow (↑) is the site of the single amino acid change (G224V) in the ***pmp-su1*** strain. **A#6** is the *LCI15* mutant generated by an *LCI15*-targeting amiRNA transformation in a *pmp1* background. The 3 green asterisks (***) are placed underneath the three-amino acid deletion that is predicted in the translated protein. **Cas9-27, Cas9-28, and Cas9-19** are the 3 mutants generated by CRISPR/Cas9-mediated, *LCI15*-targeting sgRNA editing. These 3 strains are predicted to encode heavily modified and truncated proteins. The residue positions from which the deletions (and subsequent frame-shift mutations) in their genomic sequences result in amino acid changes are highlighted by purple boxes (▬). The two recognizable domains in the predicted wild-type protein – a cobalamin synthesis protein CobW-like C-terminal domain (▬) and a WW (two tryptophan residues) domain (▬) – are indicated by a colored bar above the protein sequences.

The protein has seven WD-40 tandem repeats in its C-terminus. A WD-40 domain consists of 40 amino acid repeats that terminate in a tryptophan (W) and an aspartic acid (D) residue and has a doughnut shape with a "hole" that often mediates interactions with other proteins (Schapira *et al.*, 2017; van der Voorn and Ploegh, 1992).

Multiple WD-40 tandem repeats are typical for the WD-repeat family of proteins. This family comprises proteins that, in eukaryotes, are involved in diverse cellular machineries responsible for RNA processing, transcription initiation and regulation, signal transduction, apoptosis etc., and are especially widespread in transcriptional processes and chromatin modification (Hu *et al.*, 2017; van Nocker and Ludwig, 2003; Xu and Min, 2011). Among the ten most commonly observed in eukaryotic proteomes, WD-40 (or WD-repeat) domains within a protein play a prominent role in facilitating protein-protein, protein-peptide, and protein-DNA interactions, and structural studies have revealed that they can act as scaffolds for the assembly of larger protein complexes and that they are ideal for providing a central platform for interaction networks within the cell (Stirnimann *et al.*, 2010). WD-40 domains have not shown enzymatic activity in various studies, further supporting the notion that, rather than play a direct role in the protein's function, they play an important structural role that enables the function of a particular protein (Neer *et al.*, 1994; Stirnimann *et al.*, 2010). Examples of WD-40-repeat-containing protein complexes are G proteins and TAFII (TBP-associated factor) proteins, which are part of the transcription initiation factor TFIID protein complex (Li and Roberts, 2001; Smith *et al.*, 1999).

Németh *et al.* (1998) found that *prl1* mutations in Arabidopsis resulted in transcriptional derepression of some glucose responsive genes *and* in activation of others,

concluding that PRL1 may act as a negative regulator. The mutations also led to a simultaneous hypersensitivity to glucose. *prl1* mutant plants were of a smaller size compared to WT plants, and apart from other growth defects, developed roots that were 2- or 3-fold shorter than those in WT (Szakonyi, 2006). PRL1 was demonstrated to exhibit glucose-dependent interaction with Arabidopsis homologs of SNF1 (the yeast serine/threonine protein kinase), AKIN10 and AKIN11, in yeast two-hybrid experiments and is reportedly a potential subunit of these proteins (Bhalerao *et al.*, 1999). It was observed that *PRL1* mutants had a higher activation level of AKIN immunocomplexes and that the presence of PRL1 inhibits the kinase activity of AKIN10 and AKIN11 *in vitro*, indicating that this protein negatively regulates the Arabidopsis SNF1 homologs.

Zhang *et al.* (2014) determined that, in Arabidopsis, PRL1 is involved in primary microRNA (miRNA) transcript processing and positively regulates the accumulation of these miRNAs and small-interfering RNAs (siRNAs), therefore a decrease in the levels of these RNAs was observed in *prl1* mutants. PRL1 reportedly interacts with CDC5 (Cell Division Cycle 5), a DNA-binding protein suggested to be a transcription factor of *MIR* (miRNA encoding genes) (Palma *et al.*, 2007; Zhang *et al.*, 2013). Both these proteins interact with the DCL1 (Dicer-like 1) complex, and the PRL1-DCL1 interaction apparently is important for the processing of primary miRNAs (pri-miRNAs).

A number of proteins were identified as interactors with PRL1 via *in vitro* protein-binding assays (Salchert, 1997; Szakonyi, 2006). One of these was named PIP-L (PRL1 interacting protein-L) and was found to interact with PRL1 towards the latter's C-terminal end, downstream of its sixth WD-40 repeat. Mutants in various *PIP* (PRL1-interacting protein) genes (including *PIP-L*) were generated and crossed to *prl1* mutants

to construct double mutants; homozygous double mutants also exhibited the root elongation defect of *prl1* mutants, suggesting that the PIPs are not involved in root elongation pathways. The authors do not reveal other phenotypic effects, if any, that the *pip* mutations have on the plants. PIP-L is predicted to be localized to the chloroplast stroma (Ferro et al., 2010).

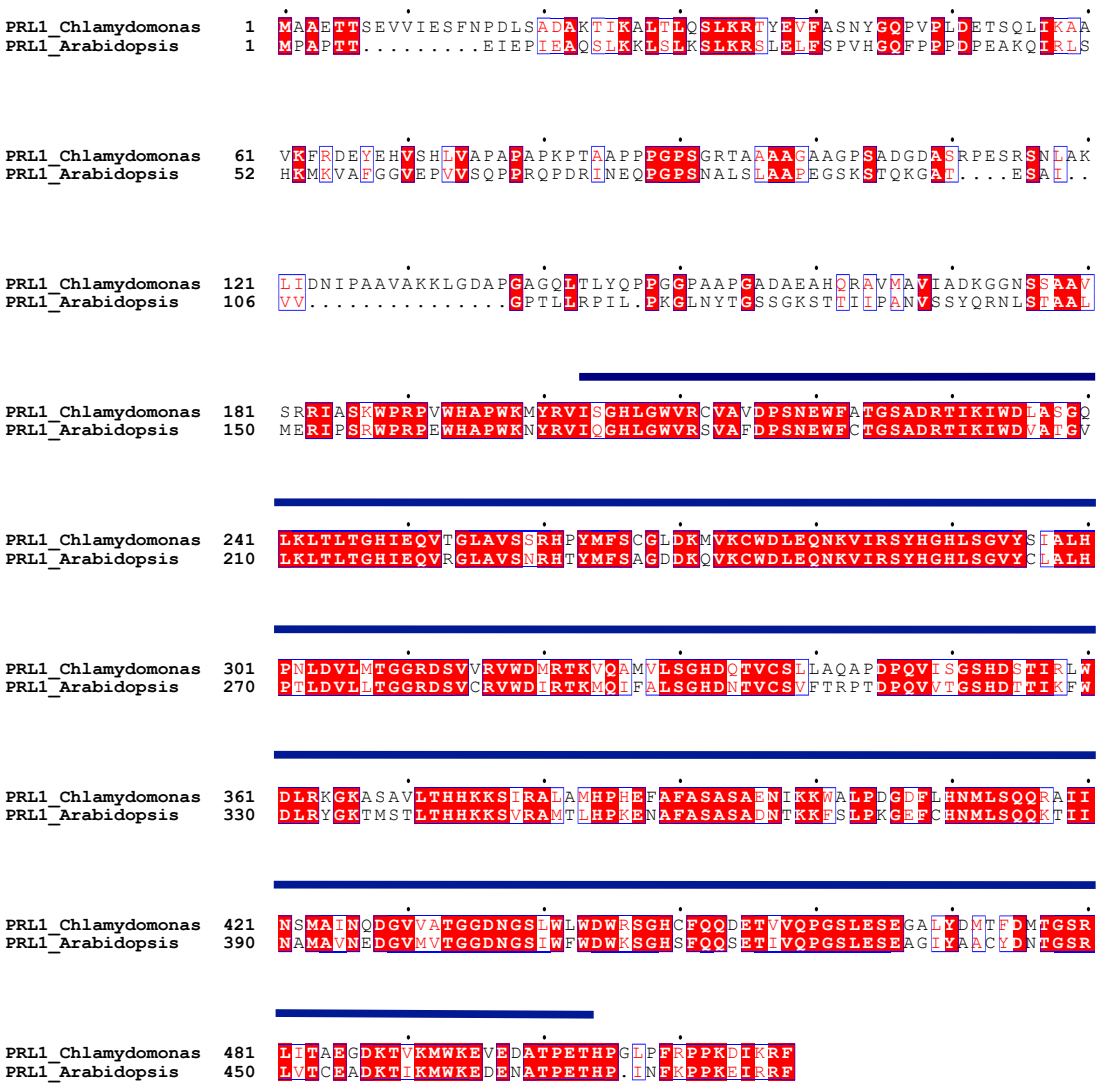


Figure 20. Sequence comparison of predicted PRL1 proteins in Chlamydomonas and in Arabidopsis. The blue bar (▬) represents the WD-40-repeat-containing domain in both proteins.

The predicted PRL1 in Chlamydomonas (Cre10.g446900) and PRL1 in Arabidopsis (At4g15900) show 56% identity and 69% similarity, and both have seven predicted WD-40 domain repeats (alignment shown in Figure 20). LCI15 in Chlamydomonas (Cre16.g685050) and PIP-L in Arabidopsis (At1g15730) show 17% identity and 29% similarity, with the latter also having a predicted CobW-like C-terminal domain but no WW domain (Figure 21).

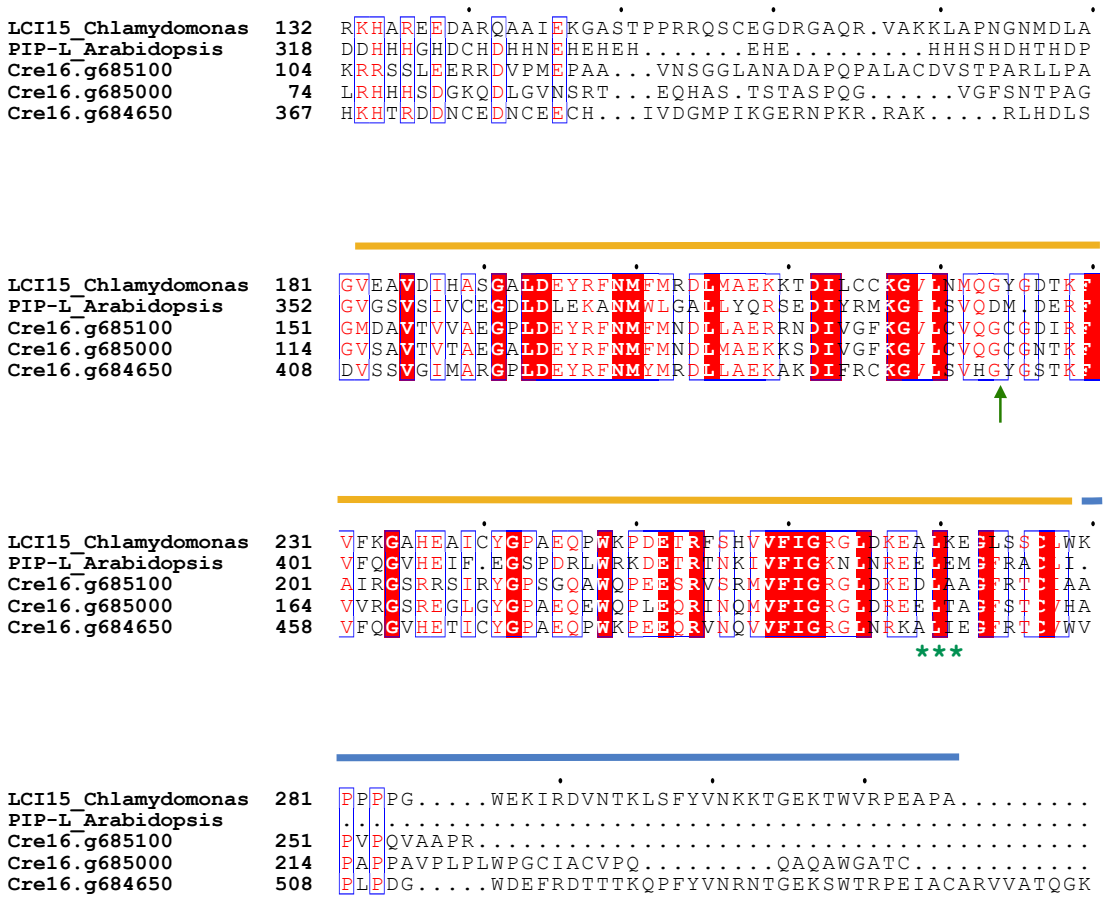


Figure 21. Sequence comparison of predicted protein sequences of LCI15 from Chlamydomonas, PRL1 interacting protein-L (PIP-L) from Arabidopsis, and three LCI15 paralogs in Chlamydomonas – Cre16.g685100, Cre16.g685000, and Cre16.g684650. The green arrow (↑) is the site of the single amino acid change (G224V) in the *pmp-su1* strain. The two recognizable domains in LCI15 – a cobalamin synthesis protein CobW-like C-terminal domain (▬) and a WW (two tryptophan residues) domain (▬) – are indicated by a colored bar above the protein sequences.

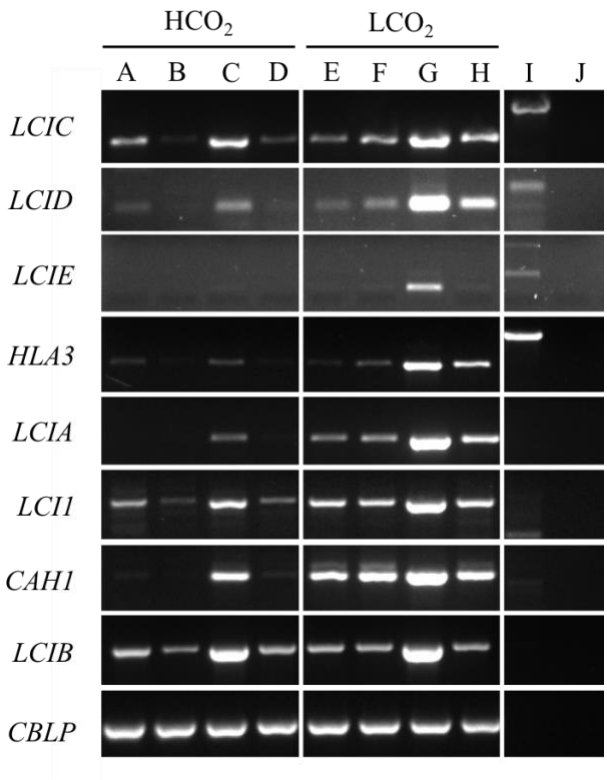
LCI15 Paralogs in Chlamydomonas

Three genes paralogous to *LCI15* are present in the same vicinity of chromosome 16 in tandem repeats: two upstream (Cre16.g684950, Cre16.g685000) and one downstream (Cre16.g685100) of *LCI15*. While the predicted translation products of all three genes have predicted CobW-like C-terminal domains, none, other than LCI15 itself, have predicted WW domains.

As stated earlier, many of the *LCI15* mutants generated are predicted to make highly truncated and modified proteins in which the CobW-like and WW-like domains are completely missing. Also, two *LCI15* mutants, *pmp-su1* and A#6, with the smallest disruptions in their predicted amino acid sequences, have an amino acid substitution and a three-residue deletion, respectively, within the CobW-like C-terminal domain (Figure 19). This suggests that this domain may be important to the protein, perhaps for its structural integrity and/or for the proper functioning of the resulting LCI15 protein.

CCM-Related Gene and Protein Expression in LCO_2-Acclimated Cells

Semi-quantitative RT-PCR analyses showed a noticeable increase in transcript abundance of CCM-related genes amplified from LCO_2-acclimated *pmp-su1* and *wt-su1* lines, as compared to those in WT and *pmp1* lines, i.e., an increase in *su1* mutant lines over non-*su1* (or wild-type *LCI15* allele) lines (Figure 22). These include *LCIB, LCIC, LCID, LCIE, HLA3, LCIA, LCI1,* and *CAH1,* with an apparent greater increase in *wt-su1* than in *pmp-su1*. Initial experiments show that *LCIC*, *LCID*, *LCI1*, *HLA3*, *LCIA*, and *CAH1* levels are higher in *wt-su1* compared to *pmp-su1*.



cDNA templates used:

A–D: HCO_2-acclimated cells
- **A**: 2137
- **B**: *pmp1*
- **C**: *wt-su1*
- **D**: *pmp-su1*

E–H: LCO_2-acclimated cells
- **E**: 2137
- **F**: *pmp1*
- **G**: *wt-su1*
- **H**: *pmp-su1*

- **I**: genomic DNA *(negative control)*
- **J**: H_2O *(negative control)*

Figure 22. Semi-quantitative RT-PCR amplification results for CCM-related genes in *LCI15* mutant and non-mutant strains (only 35-cycle PCR product gel electrophoresis pictures are shown here). Cells were acclimated to either HCO_2 or LCO_2. cDNA from equal amounts of RNA was used as template for each of the sample PCRs. Genomic DNA and water were used as controls.

As seen in Figure 23, LCI1 and LCIA immunoblots detected higher levels of these proteins in LCO_2-grown *pmp-su1* relative to WT and *pmp1*, but an even higher increase in the amounts of these proteins were observed in *wt-su1* relative to the other LCO_2-acclimated strains – in *wt-su1* versus WT, LCI1 showed an approximate 2.7-fold

increase, and LCIA showed an approximate 15-fold increase. LCIB and LCIC abundance in LCO_2-*wt-su1* cells also were observed to be higher than those in the corresponding WT cells (~1.3-fold increase).

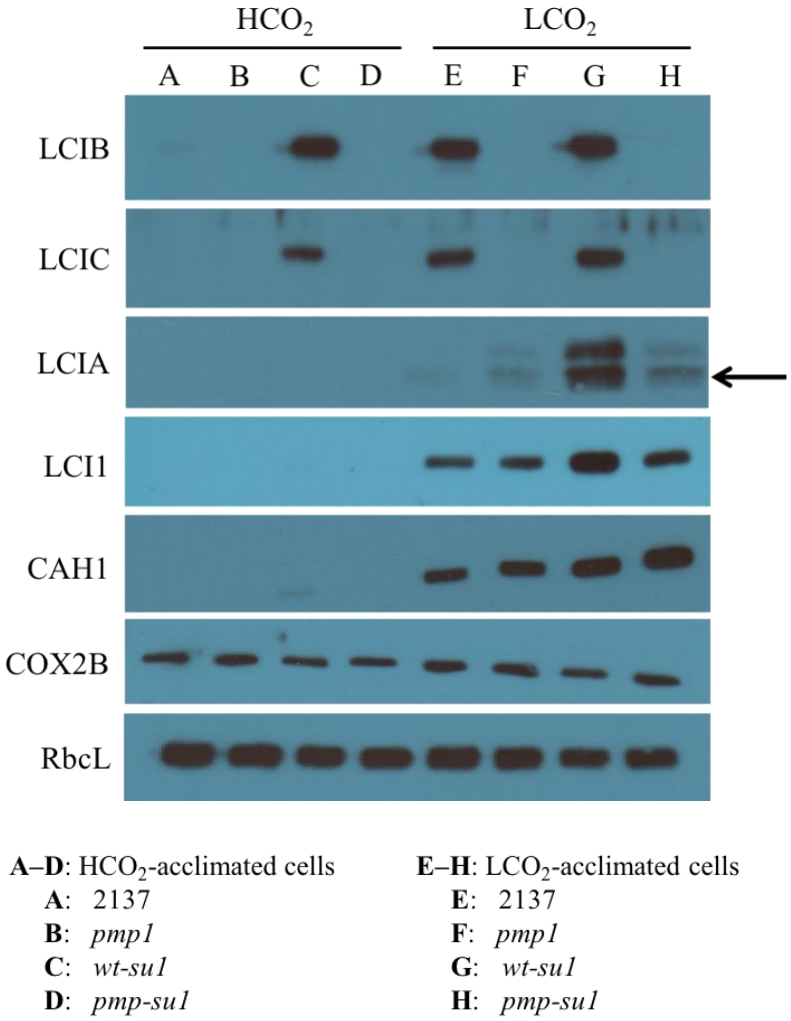


Figure 23. Western immunoblots for CCM-related proteins in *LCI15* mutant and non-mutant strains. Cells were acclimated to either HCO_2 or LCO_2. COX2B and RbcL immunoblots were used as loading controls.

CCM-Related Gene and Protein Expression in HCO_2-Acclimated Cells

A pattern similar to that observed in LCO_2-acclimated cells was seen in HCO_2-acclimated cells: higher transcript abundances for CCM genes were observed in *wt-su1* HCO_2-acclimated cells relative to those in WT cells (Figure 22).

LCIB and LCIC, which are constitutively expressed in WT cells in HCO_2 and upregulated in limiting CO_2 (Fang *et al.*, 2012; Miura *et al.*, 2004; Wang and Spalding, 2006), appeared to be upregulated at the protein level in *wt-su1* even in HCO_2-acclimated cells (Figure 23). However, for other CCM proteins tested, including LCI1, CAH1, and LCIA, no visible increase in protein under HCO_2 conditions was detected, possibly because they are not typically expressed at all in HCO_2 and are induced only in LCO_2. Since we did not have antisera for LCID and LCIE, the levels of these proteins in HCO_2-grown *su1* cells could not be surveyed, although their transcript abundance did apparently increase. If they have an expression pattern similar to that seen in LCIB and LCIC, the other members of the LCIB gene family, which exhibit upregulation in *wt-su1* HCO_2-grown cells, it would be informative as to whether it is even possible for LCID and/or LCIE to be in a potential LCIB-replacement protein complex. This would be speculative, of course, since without additional work it would not be clear if they are somehow masking the effect of LCIB.

Discussion

Our evaluation of the nature of LCI15 based on its sequence offers few clues as to how it might function in the cell, let alone how its absence might lead to photosynthesis and growth of the *LCIB* mutant Chlamydomonas cells in the normally-lethal LCO_2 conditions. The only possible clue comes from our observation that all *LCI15* mutants

identified so far as suppressors of the *LCIB* mutant *air dier* phenotype either have amino-acid changes in the CobW-like C-terminal domain of LCI15 or are truncated and lack this C-terminal domain entirely. However, these observations are limited up to now. The truncated versions of LCI15 might not even be expressed, as would be expected also for the insertional alleles of *LCI15*, so the fact that they are missing the CobW-like domain may be moot. That would leave just two mutant LCI15s we might consider likely to be expressed as translation products in which the CobW-like domain has been altered. This is too small a sample from which to draw any secure conclusions, but it is tantalizing enough to argue for future experiments.

In a similar vein, it is at least interesting that, all three identified other paralogs of LCI15 in Chlamydomonas contain the CobW-like domain but all three lack the predicted WW-like, putative protein-interaction domain. Since it is clear that none of these paralogs can compensate for the loss of LCI15, perhaps their inability to do so stems from an inability to interact with a needed protein partner via this WW-like domain. Again, this is based on very limited evidence but also is tantalizing enough to warrant additional, future exploration.

Here, we attempt to explain what the role of LCI15 is in the CCM and how it might be facilitating a replacement for LCIB in *LCIB-LCI15* mutants. One of the first things we wished to address was how the *su1*-suppressed *LCIB* mutant cells are able to circumvent the need for LCIB function in Ci uptake under air-level CO_2. Our experiments and observations point to a couple of viable possibilities. Transcript and protein abundance for a variety of CCM-related genes are higher in *LCI15* mutant lines than in non-mutant lines, which suggests that an increase in overall CCM activity might

be involved in compensating for the lack of LCIB. Also, the increased abundance of *LCID* and *LCIE* transcript abundance in *LCI15* mutant lines is consistent with the possibility of an alternative LCIB-family complex functioning to compensate for the lack of a functional LCIB/LCIC complex.

RT-PCR amplification and protein immunoblotting for some limiting-CO_2-inducible genes transcripts and their translation products revealed an increase in both transcript and protein abundance in the suppressor mutants relative to those in WT cells grown in LCO_2. There was an increase in mRNA abundance observed in *su1* (*wt-su1* and *pmp-su1*) cells over non-*su1* mutant (WT and *pmp1*) cells for the *LCIB* gene family (*LCIB*, *LCIC*, *LCID*, and *LCIE*); for *HLA3* and *LCIA*, whose translation products are associated with HCO_3^- uptake; for *CAH1*, which encodes a periplasmic carbonic anhydrase; and for *LCI1*, which encodes a plasma membrane-localized putative Ci transporter. Any, some combination of, or all of these CCM components, if present in greater levels might conceivably at least help to compensate for a lack of LCIB function in air-level CO_2.

Northern blots performed earlier (Duanmu and Spalding, 2011) also detected an increased mRNA abundance for *HLA3* in LCO_2-grown *pmp-su1* cells in comparison to that in *pmp1* cells. In addition, there was visibly more *HLA3* transcript in *wt-su1* than in WT (CC-620); this is consistent with the observations in our experiments. In the previous analyses, it was concluded that there was *retention* in transcript abundance from *pmp-su1* to *wt-su1*, while an *increase* was observed in *HLA3* (and other genes) from the RT-PCR results in this report. This difference in interpretation can perhaps be explained by the difference not being large enough to detect as an absolute value/quantity on a Northern

blot but being obvious after many cycles of amplification through RT-PCR. Nevertheless, the apparent increase in transcript abundance in the *su1* suppressor mutants suggests that LCI15 might be involved in negative regulation of these genes in LCO_2 levels, and, since the change is seen in mRNA quantity, the regulation might be at the transcriptional level. It should be noted that the differences in transcript abundance between WT, *pmp1*, *wt-su1*, and *pmp-su1* that were seen in these RT-PCR results here were not all detected in the CCM genes included in Duanmu *et al.*'s (2011) Northern analyses, which included *CAH1, LCIC, LCIA, LCI1,* and *CCP1/2*.

In HCO_2-acclimated cells, higher mRNA accumulation and protein accumulation were observed for *LCIB* and *LCIC* in *wt-su1* relative to those in WT. There were detectable levels of CAH1 and LCI1 proteins in HCO_2-grown *wt-su1* that were not observed in WT cells. These observations are consistent with LCI15 playing a part (either directly or indirectly) in repressing the CCM, since expression of at least some genes in the mechanism is induced or up-regulated in *LCI15* mutants.

Since all four members of the LCIB family are among those genes with apparently increased expression in *LCI15* mutants, an obvious possibility to compensate for the lack of LCIB would be a substitute LCIB/LCIC-like complex composed of one or more of the LCIB family proteins. From our work, it is clear that LCIB itself is still completely absent (i.e., the *pmp1* point mutation has not reverted where *pmp1* is the *LCIB* mutant allele was used) and that this putative replacement complex cannot include LCIC, since, just as in *LCIB* mutants, there is no detectable LCIB or LCIC in *LCIB-LCI15* mutants. However, it is still possible that LCID and LCIE might form either homomultimers or heteromultimers that function to replace the lost LCIB/LCIC complex.

Based on the increase in their transcript abundance in LCO_2-grown *su1* cells, it is tempting to assume that there may also be increases in the LCID and/or LCIE translation products that can then form a functional complex. Unfortunately, we do not have antibodies with which to perform western immunoblots on these two proteins to see if there is any increase in their protein abundance between WT and *su1* strains.

The possibility that there is an alternate pathway that is upregulated, or that there is relief from the CO_2-related inhibition of the LCIA-associated HCO_3^- pathway cannot be ruled out based on these results.

Based on the increased expression of several CCM-related, LCO_2-inducible and LCO_2-upregulated genes in multiple *LCI15* mutants, including clearly null alleles, it appears that LCI15 is involved in some way in negative regulation of these genes and perhaps of the CCM in general. Furthermore, since in the absence of LCI15, or of LCI15 function in some cases, there is an increase in these CCM genes' mRNA abundance, the impact of the LCI15-associated regulation might be occurring at the transcriptional or post-transcriptional level, perhaps at the level of transcript stability as described below. If LCI15 is involved in negative regulator of transcription (for example, as a repressor) or if it is involved in a post-transcriptional gene silencing mechanism (for example, in the degradation of mRNA or prevention of translation of the mRNA), the presence of LCI15 will manifest itself as a decrease in mRNA levels and the absence of LCI15 will be noticeable as an increase in mRNA levels. These proposals for LCI15 function are speculative and more work needs to be to test them.

PRL1 in Arabidopsis, as discussed earlier, is reported to be important for processing of miRNAs. If LCI15/PRL1 interacting factor-L has a similar role in

Chlamydomonas and it aids in the processing of CCM-gene-targeting miRNAs, absence of a functional LCI15 might decrease turnover of these transcripts and thus cause an increased transcript abundance of those CCM genes. However, more work will need to be done to determine what exactly LCI15 does in the CCM or how it might be negatively regulating CCM genes. It will also be interesting to see if PRL1 in Arabidopsis and in Chlamydomonas have similar functions, and what the relation between LCI15 and PRL1 in Chlamydomonas might be.

In conclusion, the two possibilities of (i) increased expression of CCM genes compensating for lack of LCIB and (ii) increased expression of LCID and/or LCIE to form an alternate complex are not mutually exclusive and could be functioning either separately or together to compensate for the lack of LCIB.

References

Bedford, M.T., Sarbassova, D., Xu, J., Leder, P., & Yaffe, M.B. 2000. A novel pro-Arg motif recognized by WW domains. *J Biol Chem* 275:10359-10369.

Bertrand, E.M., Saito, M.A., Jeon, Y.J., & Neilan, B.A. 2011. Vitamin B12 biosynthesis gene diversity in the Ross Sea: the identification of a new group of putative polar B12 biosynthesizers. 13:1285-1298.

Bhalerao, R.P., Salchert, K., Bako, L., Okresz, L., Szabados, L., Muranaka, T., Machida, Y., Schell, J., & Koncz, C. 1999. Regulatory interaction of PRL1 WD protein with *Arabidopsis* SNF1-like protein kinases. *Proc Natl Acad Sci U S A* 96:5322-5327.

Bruce, M.C., Kanelis, V., Fouladkou, F., Debonneville, A., Staub, O., & Rotin, D. 2008. Regulation of Nedd4-2 self-ubiquitination and stability by a PY motif located within its HECT-domain. *Biochem J* 415:155-163.

Buchanan, B.B., Gruissem, W., & Jones, R.L. 2015. Biochemistry and Molecular Biology of Plants, 2nd ed. John Wiley and Sons.

Bunai, K., Okubo, H., Hano, K., Inoue, K., Kito, Y., Saigo, C., Shibata, T., & Takeuchi, T. 2017. TMEM207 hinders the tumour suppressor function of WWOX in oral squamous cell carcinoma. *J Cell Mol Med*.

Dodd, A.N., Borland, A.M., Haslam, R.P., Griffiths, H., & Maxwell, K. 2002. Crassulacean acid metabolism: plastic, fantastic. *Journal of Experimental Botany* 53:569-580.

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 Ci affinity in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 106:5990-5995.

Duanmu, D., & Spalding, M.H. 2011. Insertional suppressors of *Chlamydomonas reinhardtii* that restore growth of air-dier *lcib* mutants in low CO_2. *Photosynth Res* 109:123-132.

Fang, H., Kang, J., & Zhang, D. 2017. Microbial production of vitamin B12: a review and future perspectives. *Microbial Cell Factories* 16:15.

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.

Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S., Kieffer-Jaquinod, S., Bruley, C., Garin, J., Joyard, J., Masselon, C., & Rolland, N. 2010. AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol Cell Proteomics* 9:1063-1084.

Furbank, R.T., Hatch, M.D., & Jenkins, C.L.D. 2000. C4 Photosynthesis: Mechanism and Regulation. Pp. 435-457 in: Leegood, R.C., Sharkey, T.D., & von Caemmerer, S., (eds), *Photosynthesis: Physiology and Metabolism*. Springer Netherlands, Dordrecht. p 435-457.

Hashimoto, Y., Nishiyama, M., Horinouchi, S., & Beppu, T. 1994. Nitrile hydratase gene from *Rhodococcus sp*. N-774 requirement for its downstream region for efficient expression. *Biosci Biotechnol Biochem* 58:1859-1865.

Hu, X.J., Li, T., Wang, Y., Xiong, Y., Wu, X.H., Zhang, D.L., Ye, Z.Q., & Wu, Y.D. 2017. Prokaryotic and Highly-Repetitive WD40 Proteins: A Systematic Study. *Sci Rep* 7:10585.

Huang, X., Beullens, M., Zhang, J., Zhou, Y., Nicolaescu, E., Lesage, B., Hu, Q., Wu, J., Bollen, M., & Shi, Y. 2009. Structure and function of the two tandem WW domains of the pre-mRNA splicing factor FBP21 (formin-binding protein 21). *J Biol Chem* 284:25375-25387.

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 U S A* 113:14716-14721.

Jokel, M., Kosourov, S., Battchikova, N., Tsygankov, A.A., Aro, E.M., & Allahverdiyeva, Y. 2015. *Chlamydomonas* Flavodiiron Proteins Facilitate Acclimation to Anoxia During Sulfur Deprivation. *Plant Cell Physiol* 56:1598-1607.

Kay, B.K., Williamson, M.P., & Sudol, M. 2000. The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J* 14:231-241.

Li, D., & Roberts, R. 2001. WD-repeat proteins: structure characteristics, biological function, and their involvement in human diseases. *Cell Mol Life Sci* 58:2085-2097.

Li, Y., Yang, F.L., Zhu, C.F., & Tang, L.M. 2017. Effect and mechanism of RNAi targeting WWTR1 on biological activity of gastric cancer cells SGC7901. *Mol Med Rep*.

Lu, P.-J., Zhou, X.Z., Shen, M., & Lu, K.P. 1999. Function of WW Domains as Phosphoserine- or Phosphothreonine-Binding Modules. *Science* 283:1325.

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.e114.

Martens, J.H., Barg, H., Warren, M., & Jahn, D. 2002. Microbial production of vitamin B12. *Applied Microbiology and Biotechnology* 58:275-285.

Matsuo, T., & Ishiura, M. 2010. New insights into the circadian clock in *Chlamydomonas*. *Int Rev Cell Mol Biol* 280:281-314.

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* 135:1595-1607.

Neer, E.J., Schmidt, C.J., Nambudripad, R., & Smith, T.F. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371:297-300.

Nemeth, K., Salchert, K., Putnoky, P., Bhalerao, R., Koncz-Kalman, Z., Stankovic-Stangeland, B., Bako, L., Mathur, J., Okresz, L., Stabel, S., Geigenberger, P., Stitt, M., Redei, G.P., Schell, J., & Koncz, C. 1998. Pleiotropic control of glucose and hormone responses by PRL1, a nuclear WD protein, in *Arabidopsis*. *Genes Dev* 12:3059-3073.

Palma, K., Zhao, Q., Cheng, Y.T., Bi, D., Monaghan, J., Cheng, W., Zhang, Y., & Li, X. 2007. Regulation of plant innate immunity by three proteins in a complex conserved across the plant and animal kingdoms. *Genes Dev* 21:1484-1493.

Robert, X., & Gouet, P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42:W320-324.

Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., & Gelfand, M.S. 2003. Comparative genomics of the vitamin B12 metabolism and regulation in prokaryotes. *J Biol Chem* 278:41148-41159.

Sage, R.F. 2004. The evolution of C4 photosynthesis. *New Phytologist* 161:341--370.

Salchert, K. 1997. Untersuchungen zur Funktion von PRL1 durch Identifizierung und Charakterisierung von PRL1-interagierenden-Proteinen.

Sambrook, J., and Russell, R.W. 2001. Molecular Cloning: A Laboratory Manual, 3 ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Schapira, M., Tyers, M., Torrent, M., & Arrowsmith, C.H. 2017. WD40 repeat domain proteins: a novel target class? *Nat Rev Drug Discov* 16:773-786.

Smerdon, S.J., & Yaffe, M.B. 2010. Chapter 72 - Recognition of Phospho-Serine/Threonine Phosphorylated Proteins by Phospho-Serine/Threonine-Binding Domains A2 - Bradshaw, Ralph A. Pp. 539-550 in: Dennis, E.A., (ed), *Handbook of Cell Signaling (Second Edition)*. Academic Press, San Diego. p 539-550.

Smith, T.F., Gaitatzes, C., Saxena, K., & Neer, E.J. 1999. The WD repeat: a common architecture for diverse functions. *Trends Biochem Sci* 24:181-185.

Spalding, M.H., Spreitzer, R.J., & Ogren, W.L. 1983. Reduced Inorganic Carbon Transport in a CO_2-Requiring Mutant of Chlamydomonas reinhardii. *Plant Physiol* 73:273-276.

Spreitzer, R.J., & Mets, L. 1981. Photosynthesis-deficient Mutants of *Chlamydomonas reinhardii* with Associated Light-sensitive Phenotypes. *Plant Physiol* 67:565-569.

Stirnimann, C.U., Petsalaki, E., Russell, R.B., & Müller, C.W. 2010. WD40 proteins propel cellular networks. *Trends Biochem Sci* 35:565-574.

Sudol, M., & Hunter, T. 2000. NeW wrinkles for an old domain. *Cell* 103:1001-1004.

Szakonyi, D. 2006. Genetic dissection of regulatory domains and signalling interactions of PRL1 WD-protein in *Arabidopsis*. In. Universität zu Köln.

van der Voorn, L., & Ploegh, H.L. 1992. The WD-40 repeat. *FEBS Letters* 307:131-134.

van Nocker, S., & Ludwig, P. 2003. The WD-repeat protein superfamily in *Arabidopsis*: conservation and divergence in structure and function. *BMC Genomics* 4:50.

Vance, P., & Spalding, M.H. 2005. Growth, photosynthesis, and gene expression in *Chlamydomonas* over a range of CO_2 concentrations and CO_2/O_2 ratios: CO_2 regulates multiple acclimation states. *Canadian Journal of Botany* 83:796-809.

Wang, Y., & Spalding, M.H. 2006. An inorganic carbon transport system responsible for acclimation specific to air levels of CO_2 in *Chlamydomonas reinhardtii*. *Proc Natl Acad Sci U S A* 103:10110-10115.

Wang, Y., & Spalding, M.H. 2014a. 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.

Wang, Y., & Spalding, M.H. 2014b. LCIB in the Chlamydomonas CO_2-concentrating mechanism. *Photosynth Res* 121:185-192.

Winck, F.V., Páez Melo, D.O., & González Barrios, A.F. 2013. Carbon acquisition and accumulation in microalgae *Chlamydomonas*: Insights from "omics" approaches. *Journal of Proteomics* 94:207-218.

Xu, C., & Min, J. 2011. Structure and function of WD40 domain proteins. *Protein Cell* 2:202-214.

Yamano, T., Miura, K., & Fukuzawa, H. 2008. Expression analysis of genes associated with the induction of the carbon-concentrating mechanism in *Chlamydomonas reinhardtii*. *Plant Physiol* 147:340-354.

Yamano, T., Tsujikawa, T., Hatano, K., Ozawa, S., 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.

Zhang, S., Liu, Y., & Yu, B. 2014. PRL1, an RNA-binding protein, positively regulates the accumulation of miRNAs and siRNAs in *Arabidopsis*. *PLoS Genet* 10:e1004841.

Zhang, S., Xie, M., Ren, G., & Yu, B. 2013. CDC5, a DNA binding protein, positively regulates posttranscriptional processing and/or transcription of primary microRNA transcripts. *Proc Natl Acad Sci U S A* 110:17588-17593.

CHAPTER 4. GENERAL SUMMARY

Many photosynthesizing organisms have evolutionarily developed ways to efficiently sequester atmospheric CO_2 around Rubisco, the enzyme that establishes the initiation of the Calvin cycle of photosynthesis. These include some terrestrial plants that employ the C4 or the Crassulacean acid metabolism (CAM) pathways, and aquatic organisms such as algae and cyanobacteria that have developed CO_2-concentrating mechanisms (CCMs). Understanding these systems and their components has important and potentially useful implications in both basic and applied research areas. Our work focusses on the CCM in the green alga, *Chlamydomonas reinhardtii*, and on understanding the changes it undergoes in its physiology, metabolism, and gene expression patterns as it acclimates to various levels of ambient CO_2. LCIB (Low CO_2-Inducible protein B), a protein that is upregulated in limiting CO_2 conditions, plays a major, sometimes indispensable, role in the active uptake of CO_2 and in preventing the loss of intracellular CO_2. Here, we sought to identify the *SU1* (*suppressor* 1) locus, which, when mutated, leads to a compensation for the lack of the LCIB. The intent was to better comprehend not only the exact function and mechanism of LCIB, but also what other protein(s) appears to be able to functionally replace it.

In Chapter 2, we discussed the methods employed to identify candidate *SU1* loci and to then confirm one of them as being in the gene of interest. In Chapter 3, our experiments were focused on gaining some insight into how the loss of LCIB is counteracted by the loss of LCI15 (the protein encoded by the *SU1* allele), and into the effects (if any) that LCI15 has on other CCM-related genes. We found that LCI15

appears to affect transcript and protein abundance of some CCM genes and discuss possible ways in which this might result in supplanting the role of LCIB.

The deep sequencing approach that we adopted to identify the location of the *su1* mutation successfully narrowed down the region of interest in the genome. This technique of sequencing pooled DNA from a collection of recombinant and independent progeny, all of which either do or do not carry the mutation of interest, is especially advantageous when working with a phenotype that is obvious and can be easily screened and/or selected. While we generated almost 500 independent, recombinant progeny with the intention of mapping the *SU1* locus as finely as possible, it might be useful in some situations to gather fewer progeny, perform "rough" mapping, and then scour the relevant region to assemble a list of candidate genes that can reasonably be expected to affect the phenotype being studied. The frequency of markers (such as SNPs) between the polymorphic strain being used and the reference strain(s) is also important when applying this method. For example, a low frequency of SNP differences in the as-yet-unmapped region of the mutation will cause the SNP ratio to be low if the sequencing coverage is low *and* there are SNPs undetected due to sequencing errors in some reads; the low SNP frequency factor is critical here because more SNPs would offset the problems stated above. Using a polymorphic strain with a fairly high average marker density will decrease the likelihood of facing such an issue.

If there is only a single mutation whose locus is unknown, or if the phenotype of the mutation is difficult to screen or select for, an antibiotic-resistant polymorphic strain can prove extremely useful as selection for recombinant progeny is then possible. Transforming the polymorphic strain in order to introduce an antibiotic resistance

cassette in it will nevertheless be a risky proposition since it can lead to multiple disruptions anywhere in the genome.

By crossing our mutant of interest with a polymorphic strain, generating a population of independent, recombinant progeny with a similar genotype in the locus of the mutation across all the progeny, deep sequencing their genomic DNA, and then using a SNP marker library to calculate SNP densities in the results of the deep sequencing, we were able to successfully identify the mutated gene. High-throughput sequencing – now a time- and cost-efficient technology – can be used to create marker maps of lines that have naturally-occurring genetic variation between them, and also to perform deep sequencing as was done in this project. This approach is particularly practical in identifying genes mutated with minimal disruption in forward genetic screens.

Successful complementation of the *su1* mutation was unworkable, possibly because we were seeking to revert a positive to a negative growth phenotype. Further work needs to be done to confirm what the function of LCI15 is and to understand how the absence of the protein suppresses the *air dier* phenotype and could also help answer the question of why complementing the *LCI15* mutation proved to be challenging.
A real-time quantitative PCR analysis on the wild-type strain and *LCIB*, *LCI15*, and *LCIB-LCI15* mutant strains could confirm the transcript abundance increases observed in the *LCI15* mutant strains over those with the *LCI15* wild-type alleles in the semi-quantitative reverse-transcriptase PCR data, and also better distinguish the *LCI15* single mutant and *LCIB-LCI15* double mutant transcript levels. Determining whether the presence of LCIB affects the amount of mRNA transcribed from CCM genes, both with and without LCI15, would also be interesting, since it could mean that LCIB is somehow

involved in retrograde signaling (Broda and Van Aken, 2018; Leister, 2012).

In addition to a real-time PCR experiment (or even instead of, given our initial semi-quantitative RT-PCR results), a transcriptome study with *LCIB* and *LCI15* single and double mutants grown in different CO_2 conditions would shed light on the extent of the relationship that LCI15 appears to have with the expression of other genes, and on how this protein's function varies in the different CO_2 acclimation states.

Western immunoblots on these same cell lines to detect LCID, LCIE, and HLA3 protein levels might help rule out certain possibilities of how the absence of LCIB is masked in *LCI15* mutants (LCID and/or LCIE forming a LCIB/C replacement complex, or the HLA3-/LCIA-associated HCO_3^- pathway being uninhibited).

By generating antiserum against LCI15 and employing a method such as a pull-down assay, potential interactors of LCI15 can be identified. This would help us understand the mechanism behind the LCI15 protein's function and could detect currently unknown proteins in the CCM that require LCI15 or that form a complex with LCI15 to function properly. Using an appropriate secondary antibody for immunofluorescence in cells grown in different CO_2 levels will reveal any distinct localization patterns that LCI15 exhibits depending on the ambient CO_2 levels.

If the *LCIA-LCIB-LCI15* triple mutant is able to sustain growth in LCO_2 conditions, we would know that any potential relief of the inhibition of the LCIA-associated pathway in LCO_2 is not the reason for suppression (since these cells are missing LCIA anyway). However, if the triple mutant cannot survive in LCO_2, the involvement of LCIA in suppression can be confirmed by reintroducing a wild-type *LCIA* gene by complementation into the triple mutant and observing restoration of growth in

LCO_2. In addition, LCIA activity in photosynthesis can be studied by comparing O_2 evolution in an *LCIB* single mutant and an *LCIA-LCIB* double mutant, both in the presence and absence of LCI15 (i.e. wild-type and mutant *LCI15* alleles).

Similar experiments combining *LCIB-LCI15* mutations with those in other CCM genes, such as *LCID, LCIE,* or *LCI1*, will give us a hint about which other protein might (or might not) be involved in the suppression activity. Using gene-editing techniques such as CRISPR/Cas9, mutant cell lines in these other genes can be generated.

Additionally, the effect, if any, of mutations in the three *LCI15*-like genes (Cre16.g684650, Cre16.g685000, and Cre16.g685100) would also be interesting to see. However, considering that their encoded proteins did not suppress the *su1* suppressor mutation itself, i.e., they did not apparently take over the function(s) of LCI15 (or were unable to do so on their own), these mutant lines might not prove to be very instructive. We do not know what PRL1 does in Chlamydomonas, so observing the growth phenotype of a *PRL1* mutant would be particularly interesting if it was similar to that of an *LCIB-LCI15* mutant; this, along with identification of proteins interacting with LCI15, might be a first step in trying to establish the mechanism behind the LCI15 protein's role and understanding any parallels that might exist with the Arabidopsis PRL1.

How the absence of LCI15 restores growth in LCO_2 conditions and what the broader function of LCI15 is can be better understood by performing additional work on these strains. This could uncover as yet unidentified proteins and/or mechanisms, possibly in regulation, in certain growth conditions, and thus lead to a better understanding of the Chlamydomonas CCM.

References

Broda, M., & Van Aken, O. 2018. Studying Retrograde Signaling in Plants. *Methods Mol Biol* 1743:73-85.

Leister, D. 2012. Retrograde signaling in plants: from simple to complex scenarios. *Front Plant Sci* 3:135.