Analysis of mutants impaired for respiratory growth in the model photosynthetic alga, *Chlamydomonas reinhardtii*

Thesis

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Andrew David Castonguay

Graduate Program in Molecular Genetics

The Ohio State University

2021

Thesis Committee

Dr. Patrice Hamel, Advisor

Dr. Juan Alfonzo

Dr. Amanda Bird

Dr. Harold Fisk

Abstract

Mitochondrial Complex I (CI), also known as NADH:ubiquinone oxidoreductase is the first and largest enzyme complex of the mitochondrial electron transport chain (ETC) and entry point for electrons from NADH. The fully assembled complex has a molecular weight of ~1 MDa and is L-shaped with a membrane arm embedded in the inner mitochondrial membrane and soluble arm protruding into the mitochondrial matrix. Together, eukaryotic CI is composed of more than 40 subunits: 14 core subunits conserved from bacteria in addition to 25-35 non-core or accessory subunits, plus 9 non-protein cofactors (1 flavin mononucleotide and 8 iron-sulfur clusters). Due to the obvious complexity of the holoenzyme, the assembly process requires proteins not included in the final complex that are collectively termed assembly or biogenesis factors. Unexpectedly, defects in CI are implicated in a number of severe human disorders including Leigh syndrome, lactic acidosis and stroke-like episodes (MELAS) syndrome, and Parkinson's disease. However, the underlying genetic defects have been identified in only 60% of patients with CI deficiency, occurring in genes encoding a CI subunit or previously identified biogenesis factor. It is generally accepted that the causative mutations occurred in genes encoding novel CI biogenesis factors in the remaining 40% of patients.

Chlamydomonas reinhardtii is an established model system for the study of mitochondrial respiration. In contrast to other model systems, *Chlamydomonas* CI mutants

are still viable but have a characteristic slow growth in the dark (SID) phenotype in respiratory conditions (dark plus acetate). In Chapter 2, with the goal of identifying novel genes encoding factors controlling mitochondrial CI biogenesis, an insertional mutagenic screen was previously performed in *Chlamydomonas*. Of more than 54,000 insertional mutants, 22 were SID, 13 of which were also reduced for CI activity and/or assembly and termed "*amc*" for assembly of mitochondrial complex I. Importantly, of these 13 *amc* mutants, 2 (*amc5* and *amc9*) were found with lesions disrupting nuclear genes encoding CI subunits, thereby validating the efficacy of this approach. Intriguingly, each of the disrupted genes had homologues in humans and were null for their respective transcripts, providing an opportunity to model effects of mutations linked to human CI deficiency.

In Chapter 3, two additional mutants, *amc1* and *amc11* were significantly decreased for CI enzymatic activity due to a block in the assembly process resulting in accumulation of a 700 kDa subcomplex. These mutants were later found to be allelic, harboring insertions in exon 2 of *Cre16.g688900*, a predicted protein coding gene not previously affiliated with CI biology. These mutants were renamed *amc1-1* and *amc1-2*, and the defective gene, *Cre16.g688900*, as *AMC1*, encoding a large protein of biased amino acid composition and lacked any conserved domains. As expected, the N-terminus of AMC1 was sufficient to target a heterologous reporter to yeast mitochondria. Notably, other large low complexity proteins have been implicated in chloroplast gene expression, and subsequent investigation of mitochondrial transcripts revealed a specific decrease in the *nd4* transcript. Intriguingly, loss of ND4 results in the same assembly intermediate observed in *amc1-1* and *amc1-2* mutants. Taken together, these results are consistent with a function of AMC1 in

mitochondrial gene expression, specifically in expression of the *nd4* subunit, but the precise stage remains to be experimentally determined.

A fifth mutant obtained from this screen, *amc12,* was deficient in multiple respiratory enzyme complexes of the ETC but subsequently found to have a second, independently segregating mutation contributing to the SID phenotype. The single mutant (hereafter *lcla1*), which was genetically linked to the insertional cassette, was isolated and characterized in Chapter 4. Unusual among the collection of *amc* mutants, *lcla1* was also severely affected for growth in high light, and instead resembled the isocitrate lyase (*icl*) mutant. This light sensitivity was exacerbated by introduction of an arginine auxotrophic marker in *lcla1* and *icl* but not control strains, further suggestive of a common biological pathway. Since lipids are increased in the *icl* mutant and high light is a stress condition known to induce plastid lipid synthesis, triacylglycerol (TAGs) content was measured in *lcla1*. In normal growth conditions, cytoplasmic TAG-containing lipid bodies were increased in *lcla1* due to a defect in lipid catabolism. Furthermore, *lcla1* was resistant to treatment with cerulenin, a plastid fatty acid synthesis inhibitor, in high light suggesting the high light sensitivity was due to lipotoxicity.

The insertional cassette in *lcla1* was mapped to *Cre07.g329861*, predicted to encode a large protein with numerous low complexity regions but without any conserved domains suggestive of a biological function. Introduction of an epitope tag facilitated detection of the encoded isoforms at the expected sizes and the gene renamed LCLA1 for Low Complexity protein in Lipid Accumulation. The N-terminus of one isoform was sufficient to direct a chimeric reporter to the ER in mixotrophic conditions (light plus

acetate), the site of *de novo* peroxisome biogenesis. Shifting cultures to respiratory conditions, the reporter redistributed to puncta resembling peroxisomes. Due to these observations and the compartmentation of glyoxylate cycle and fatty acid oxidation enzymes in peroxisomes, an intriguing possibility is for a role of LCLA1 in peroxisome biogenesis. Indeed, only ~1/3 of the proteins involved in peroxisome biogenesis are conserved in *Chlamydomonas*. While the precise function of LCLA1 remains to be elucidated, here we provide support for a role in lipid metabolism, indirectly through peroxisome function.

In summary, from an insertional mutagenic screen for mitochondrial CI deficiency, two nuclear mutants encoding subunits of CI were obtained as well as two additional mutants disrupted for novel genes. The first, *AMC1*, was shown to encode a novel CI biogenesis factor controlling expression of the mitochondrial-encoded nd4 transcript. The second, *LCLA1*, encoding at least two isoforms that participate in lipid homeostasis, possibly through peroxisome biogenesis, an organelle central to lipid metabolism and acetate assimilation. Interestingly, from this screen we have identified two novel genes encoding large, low complexity proteins participating in diverse biological processes.

Dedication

To my wife who always knows when to push me to work harder – or hand me a drink.

Acknowledgments

First, I would like to sincerely thank my mentor, Dr. Patrice Hamel for allowing me to be very independent in my experimental design. I would also like to thank Dr. Birgit Alber for her advice and constructive criticisms regarding microbial central carbon metabolism and growth. Additionally, I would like to extend my gratitude to Dr. Nitya Subrahmanian, Dr. Norman Groves, and Dr. Justin North for their technical advice. Lastly, I would like to thank my committee members, Dr. Juan Alfonzo, Dr. Amanda Bird, and Dr. Harold Fisk, as well as my friends and family for their enduring support.

Vita

2011	Chemical Biology, Saint Joseph's University, Philadelphia
2011 to Present	Graduate Research Associate and/or Graduate Teaching Associate, Department of Molecular Genetics, The Ohio State University

Publications

Subrahmanian, N, **Castonguay, AD**, Fatnes, TA, Hamel, PP. (2020) *Chlamydomonas reinhardtii* as a plant model system to study mitochondrial complex I dysfunction. *Plant Direct.* 4: 1– 16; *https://doi.org/10.1002/pld3.200*

Subrahmanian, N, **Castonguay, AD**, Remacle, C, Hamel, PP. (2020) Assembly of Mitochondrial Complex I Requires the Low-Complexity Protein AMC1 in *Chlamydomonas reinhardtii*. *Genetics.* 214: 895-911; *https://doi.org/10.1534/genetics.120.303029*

Fields of Study

Major Field: Molecular Genetics

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Overview

This thesis is an exploration of central carbon metabolism in green algae. Central carbon metabolism can be defined as a complex series of enzymatic steps in which the formation and severing of carbon-carbon bonds yields both energy and building blocks for macromolecules. In plants, which include green algae, central carbon metabolism relies on the cooperation of interconnected metabolic networks compartmented in chloroplasts, mitochondria, cytosol, and peroxisomes. In the work described here, the experimental system under study is *Chlamydomonas reinhardtii*, a unicellular green alga which can utilize both atmospheric CO_2 and acetate as carbon sources for growth. However, growth on either carbon source (i.e., the generation of cellular biomass and energy in the form of ATP) first requires assimilation. Below I provide a brief description of *Chlamydomonas* as model organism and an overview of the fate of either carbon sources used by this organism to sustain growth, with particular emphasis on acetate. In addition, I discuss the role of peroxisome in the conversion of intracellular carbon molecules and the function of mitochondria in energy production.

1.2 *Chlamydomonas* as a genetically-tractable model algal

To address questions concerning algal metabolism, our laboratory utilizes *Chlamydomonas reinhardtii* (hereafter *Chlamydomonas*) as an experimental system. *Chlamydomonas* is a unicellular, green alga belonging to *Chlorophyceae*, and which can

be distinguished by its single, cup-shaped chloroplast and two long cilia (historically referred to as flagella) [1, 2]. The *Chlamydomona*s genome has been fully sequenced and largely annotated [2, 3]. In early drafts, gene annotation was based largely on expressed sequence tags (ESTs) and sequence homology; then a subset were experimentally validated and used to train algorithms for *ab initio* gene prediction [4]. These *Chlamydomonas*-specific prediction programs facilitated gene discovery initially. More recently, high throughput RNA-seq transcriptomic data with deeper coverage was used to validate earlier versions of predicted gene models. Updated drafts of the *Chlamydomonas* genome are released every few years and are available through the Joint Genome Institute (JGI, https://phytozome.jgi.doe.gov/pz/portal.html) [2]. Due to ease of approach, current gene annotations are largely derived from sequence homology to previously characterized genes in other model systems.

Another important feature is that *Chlamydomonas* is genetically tractable, reproducing either asexually or sexually yielding meiotically recombined progeny [5]. In *Chlamydomonas*, there are two opposite mating types denoted mt^+ and mt^-, analogous to *Saccharomyces cerevisiae*. In the laboratory, gametogenesis is induced by nutrient starvation, specifically nitrogen deprivation, and mixing of opposite mating type results in formation of a diploid zygote (zygospore) [5]. A small portion of fused gametes do not form zygospores but instead yield vegetative diploid cells. Together these features enable classical (Mendelian) genetic approaches to observe either phenotypic segregation of markers from crosses or complementation testing of alleles in vegetative diploids.

Although *Chlamydomonas* has both haploid and diploid phases of its life cycle, the primary stage is haploid, allowing recessive mutations to be immediately observed phenotypically [6]. Taking advantage of this unique feature among plant model systems, several laboratories sought to generate mutant collections selecting for a variety of phenotypes including abnormal swimming, photosynthetic deficiency, respiratory deficiency, high oil production, *etc.* [7-10]. More recently, a collaborative effort led by Dr. Martin Jonikas yielded an indexed *Chlamydomonas* insertional mutant library with approximately 70% coverage of the genome [11]. This has become an invaluable resource for the *Chlamydomonas* community, analogous to the T-DNA insertional mutant collection for *Arabidopsis thaliana* or knockout library of nonessential genes for *S. cerevisiae*. However, construction of this insertional mutant library selected against mutants defective for survival in the dark such as in pathways involved in acetate assimilation or mitochondrial respiration. These pathways are required for survival in the incubation conditions used to recover the insertional mutants of this screen.

1.3 *Chlamydomonas* as a system to explore central carbon metabolism

Chlamydomonas is ideally suited to explore the metabolic routes for energy production due to versatile culture in photo-autotrophic (light + CO_2), heterotrophic (dark + acetate), or mixotrophic (light + CO_2 + acetate) conditions. [1, 12, 13]. As in land plants, *Chlamydomonas* contains four metabolically active compartments separated by at least one membrane (chloroplast, cytosol, peroxisome, and mitochondria) [1, 12-14]. The metabolic

flexibility of *Chlamydomonas* enables the dissection of the individual contribution of photosynthesis and respiration, two processes that are required for photo-autotrophic and heterotrophic growth, respectively. Indeed, *Chlamydomonas* mutants deficient for photosynthesis can survive via heterotrophic growth and reciprocally mutants deficient for respiration are able to grow photo-autotrophically. Photo-autotrophic growth requires the operation of photosynthesis, a process taking place in the chloroplast. At the thylakoid membrane, energy in the form of light is captured to produce NADPH and ATP via a series of chemical reactions referred to as photophosphorylation. NADH and ATP serve for CO_2 fixation via RuBisCO in the chloroplast stroma and the fixed CO_2 is further used to produce all the cellular constituents. Heterotrophic growth is dependent upon respiration, a process in the mitochondria that requires the oxidation of acetyl-CoA into CO_2 via the activity of the tricarboxylic acid (TCA) cycle. Reducing power in the form of NADH produced by the TCA cycle is used in a series of chemical reactions referred to as oxidative phosphorylation to generate ATP. Intermediates generated by the TCA cycle can be used to produce other cellular constituents for growth and division. However, metabolic pathways supporting growth in the dark also require the peroxisome for acetate assimilation via operation of the glyoxylate cycle.

The glyoxylate cycle functions to produce succinate from two molecules of acetyl-CoA. The reactions of the glyoxylate cycle are performed by isocitrate lyase (ICL), malate synthase (MAS1), malate dehydrogenase (MDH1), citrate synthase (CIS2), and aconitase (ACH1) [15]. Except for ICL, each of these enzymes are peroxisome localized in

Chlamydomonas. Notably, acetyl-CoA is incorporated during 1) malate synthesis from glyoxylate via MAS1 and 2) citrate synthesis from oxaloacetate by CIS2 [15].

Acetyl-CoA is formed from acetate by acetyl-CoA synthetase using Coenzyme-A, and ATP. Several isoforms of acetyl CoA synthetase have been annotated in *Chlamydomonas* [16] but the primary route for acetate assimilation appears to be performed by the peroxisomal localized ACS3 which is functionally linked to the glyoxylate cycle [15, 17, 18]. This is supported by the observation that in the *Chlamydomonas icl* mutant, the abundance of ACS3 is considerably reduced [15, 17]. Other enzymes of the glyoxylate cycle are decreased in this mutant causing acetate assimilation to be blocked. Therefore, this mutant cannot grow heterotrophically due to the inability to assimilate acetate [17]. Aside from acetate assimilation, acetyl-CoA can also be produced through catabolism of fatty acids synthesized in the chloroplast. Catabolism of fatty acids occurs via β-oxidation and the enzymes performing these steps are expected to also be compartmented in the peroxisome due to the production of hydrogen peroxide by acyl-CoA oxidase (ACX2) [19]. However, peroxisome localization of these enzymes has not been experimentally validated in *Chlamydomonas* except for ACX2 [19]. Since multiple isoforms of acetyl-CoA synthetase are present in *Chlamydomonas*, one for each metabolically active compartment, it is expected that acetyl-CoA generated through β-oxidation would be used by the glyoxylate cycle and not enter the TCA cycle directly for the synthesis of citrate from oxaloacetate by CIS1 in the mitochondria. The latter case would require transport away from the site of synthesis, the peroxisome, to the mitochondria but this has not been investigated in *Chlamydomonas*.

Mitochondria produce the majority of cellular energy in the form of ATP through mitochondrial respiration, also referred to as oxidative phosphorylation. Oxidative phosphorylation converts an electrochemical gradient in the form of protons across the inner mitochondrial membrane into chemical bond energy in the form of ATP. There are four multimeric enzyme complexes of the mitochondrial electron transport chain (ETC) and the fifth complex, ATP-synthase, that participate in this process. Three of these enzyme complexes, Complex I (NADH:ubiquinone oxidoreductase), Complex III (ubiquinone:cytochrome *c* oxidoreductase), and Complex IV (cytochrome *c* oxidase) contribute to the proton gradient. However, reductant in the form of NADH is required by mitochondrial CI (or alternative type-II NADH dehydrogenases) to help establish the electrochemical gradient. Reductant is provided by the TCA cycle which is entirely compartmented in the mitochondrial matrix (or matrix facing side of inner mitochondrial membrane as is the case for succinate dehydrogenase/CII). The primary purposes of these reactions are to 1) recycle carbon skeletons for synthesis of cofactors, amino acids, nucleotides *etc.* and 2) to produce reductant in the form of NADH. The TCA cycle enzymes: isocitrate dehydrogenase, α-ketoglutarate dehydrogenase, and malate dehydrogenase all produce NADH whereas succinate dehydrogenase (C II) contributes electrons directly to the ETC from succinate in its conversion to fumarate, another TCA cycle metabolite.

The route for uptake of acetate from the extracellular medium by *Chlamydomonas* is presently unknown, although transport is assumed to occur via diffusion through unknown protein channel(s) as acetate cannot pass through phospholipid bilayers

passively. Whereas the route for acetate import across the plasma membrane is unknown, five genes encoding proteins of the GPR1/FUN34/YaaH (GFY) family, known to facilitate acetate uptake/transport in other model systems, were found to be expressed in response to acetate supplementation in *Chlamydomonas* [20]. However, the identified GFY proteins all localized to microbodies [20] – not peroxisomes, mitochondria, or the chloroplast, compartments with known roles in acetate assimilation [16, 21]. Although the expression pattern could be growth condition dependent, the GFY proteins identified thus far do not compose the complete set of channels responsible for intracellular acetate economy so further research in this area is needed.

Central carbon metabolism in *Chlamydomonas* is interconnected between the chloroplast, cytosol, peroxisome, and mitochondria. Carbon can be assimilated either from CO_2 fixed through photosynthesis performed in the chloroplast or by acetate assimilation predominantly performed by the peroxisome. Additionally, fatty acids produced by the chloroplast can be converted into acetyl-CoA via β-oxidation in the peroxisome. Intermediary metabolites produced in this organelle can be shared with the mitochondria, the location of the TCA cycle. In chapter 4, I describe the *lcla1* mutant affected for heterotrophic, mixotrophic and phototrophic growth and provide evidence that a primary defect in peroxisome biogenesis/function might account for all the observed growth phenotypes.

Reactions of the TCA cycle produce the reductant needed for establishing the electrochemical gradient across the mitochondrial inner membrane which is ultimately required for ATP production via oxidative phosphorylation. Loss of enzymes participating

in any of these processes or regulating their function (such as through biogenesis) result in a defect in respiratory growth. To date, many mutants have been isolated in *Chlamydomonas* based on defective growth in respiratory conditions, mainly affecting mitochondrial ETC complexes I, III, and IV, but also for ICL [8, 17, 22-29]. In chapter 2, I describe several mutants deficient for respiratory growth due defects in the assembly of CI, and in chapter 3, the identification of a novel CI assembly factor.

Chapter 2. Isolation of CI mutants in the model alga *Chlamydomonas reinhardtii* to model human mitochondrial complex I dysfunction

This chapter has been previously published:

Subrahmanian, N, **Castonguay, AD**, Fatnes, TA, Hamel, PP. **(2020)** *Chlamydomonas reinhardtii* as a plant model system to study mitochondrial complex I dysfunction. *Plant Direct.* **4**: 1– 16; *https://doi.org/10.1002/pld3.200*

Contributions:

Castonguay, AD: Figure 6C, 7E-F, 10C, 11C-D, 13C, 15C, and editing/revision process.

2.1 Abstract

Mitochondrial complex I, a proton-pumping NADH: ubiquinone oxidoreductase, is required for oxidative phosphorylation. However, the contribution of several human mutations to complex I deficiency is poorly understood. The unicellular alga *Chlamydomonas reinhardtii* was utilized to study complex I as, unlike in mammals, mutants with complete loss of the holoenzyme are viable. From a forward genetic screen for complex I-deficient insertional mutants, six mutants exhibiting complex I deficiency with assembly defects were isolated. *Chlamydomonas* mutants isolated from our screens, lacking the subunits NDUFV2 and NDUFB10, were used to reconstruct and analyze the effect of two human mutations in these subunit-encoding genes. The K209R substitution in NDUFV2, reported in Parkinson's Disease patients, did not significantly affect the enzyme activity or assembly. The C107S substitution in the NDUFB10 subunit, reported in a case of fatal infantile cardiomyopathy, is part of a conserved C-$(X)_{11}$-C motif. The cysteine substitutions, at either one or both positions, still allowed low levels of holoenzyme formation, indicating that this motif is crucial for complex I function but not strictly essential for assembly. We show that the algal mutants provide a simple and useful platform to delineate the consequences of patient mutations on complex I function.

2.2 Introduction

Mitochondrial oxidative phosphorylation (OXPHOS) involves four major membrane-bound complexes (I, II, III, and IV) mediating electron transfer from the

substrates, NADH or succinate, to the terminal electron acceptor O_2 [30]. In concert with their oxidoreductase activities, complexes I, III, and IV also translocate protons across the mitochondrial inner membrane, thereby establishing the proton gradient necessary for complex V (F_1F_0 ATP synthase) to generate ATP on the matrix side [31].

With over 40 nucleus- and mitochondria-encoded subunits, mitochondrial complex I is a type-I NADH dehydrogenase [32, 33] and the largest respiratory complex in the mitochondrial inner membrane [34, 35]. Among the ~40 subunits common to all eukaryotic complexes I [36], only 14 orthologs make up the bacterial enzyme and are therefore considered to be the "core" subunits as they constitute the minimal requirement for enzymatic activity [37]. This highly conserved core is composed of seven catalytic subunits often encoded by the nuclear genome, binding the prosthetic groups (one FMN and eight iron-sulfur (Fe-S) clusters) required for oxidation of NADH, plus seven hydrophobic subunits (ND subunits) generally encoded in the mitochondrial genomes of eukaryotes [38]. The roles of the 24 noncore subunits conserved in all eukaryotic lineages, also referred to as "accessory subunits," are largely unknown. It is proposed that they have supportive roles in stabilizing the complex and/or regulating its activity [39, 40].

Complex I biogenesis in eukaryotes is a complicated process dependent upon the coordinated expression of the nuclear and mitochondrial genomes [41]. This process has attracted considerable attention as ~37% of OXPHOS disorders are characterized by isolated or combined complex I deficiency [42, 43]. Since the first report of human complex I deficiency by Morgan-Hughes *et al*. in 1979 [44], pathogenic mutations have been discovered in 20 (out of 37) nuclear-encoded subunits and all seven mitochondrially

encoded subunits of complex I [43, 45-47]. These mutations have been associated with a variety of clinical symptoms including hypertrophic cardiomyopathy, Leigh syndrome, and additional neurodegenerative disorders [47-50]. At the cellular level, patient-derived fibroblasts display a diverse range of phenotypes including decreased complex I activity and assembly, increased reactive oxygen species production, mitochondrial membrane depolarization, defective ATP production, and altered mitochondrial morphology [51, 52].

Although our knowledge of complex I deficiency is broadening, the molecular mechanisms underlying the clinical symptoms remain poorly understood. There is no clear correlation between the clinical presentation and the corresponding molecular defects [51]. For instance, different mutations in the same gene may present with alternate clinical phenotypes. In addition, there is variability in the complex I deficiency observed at the tissue and organ levels for the same patient [52, 53]. As the heterogeneity of the biochemical and clinical phenotypes adds additional layers of complexity, demonstrating the pathogenicity of a molecular lesion in humans has become a real challenge. Therefore, some mutations have been assigned a "provisional" status because their contribution to the disease phenotype remains uncertain [54].

Due to the above-mentioned difficulties associated with studying complex I disorders; non-human experimental model systems have been used to dissect the molecular bases of mitochondrial complex I assembly. While bacterial systems have been previously used for reconstructing human pathogenic mutations, they lack the subunit complexity of their eukaryotic counterpart [55]. Similarly, the single-celled eukaryote *Saccharomyces cerevisiae* is an unsuitable experimental system because it lacks mitochondrial complex I

[56]. Previously, the obligate aerobic yeasts *Yarrowia lipolytica* and *Neurospora crassa* have been successfully utilized to mimic disease-associated mutations in genes encoding structural subunits and an assembly factor [57-60].

The unicellular photosynthetic alga *Chlamydomonas reinhardtii* (to be referred to as *Chlamydomonas*) has emerged as an alternative simple model system for studying mitochondrial complex I [8, 29, 38]. Firstly, the composition of complex I in *Chlamydomonas* is similar to its human counterpart [22, 61, 62]. Secondly, the nuclear and mitochondrial genomes encoding complex I subunits are amenable to manipulation [8, 63]. Thirdly, unlike mammalian organisms, complete loss of complex I is still viable due to the capacity of this alga to photosynthesize [25, 64]. In addition, alternative enzymes in the *Chlamydomonas* electron transport chain (ETC) can partially bypass the lack of complex I [65], thereby allowing respiratory growth due to which complex I mutants display a characteristic slow-growth-in-the-dark (SID) phenotype. In a previous study by our group, a forward genetic screen conducted based on the SID phenotype led to the isolation of seven nuclear mutants, *amc1* to *amc7* (for *a*ssembly of *m*itochondrial *c*omplex I) defining six distinct loci required for complex I function [8]. In this study, we report the description of *amc8* to *amc13* which were also uncovered via insertional mutagenesis. Among these mutants, the *amc5* and *amc9* mutations were mapped to nuclear genes encoding the complex I subunits NUOB10 (NDUFB10 in human) and NUO5 (NDUFV2 in human), respectively [8], proving the efficacy of our screen. We have utilized *Chlamydomonas* complex I mutants *amc5* (*nuob10/ndufb10-null*) and *amc9* (*nuo5/ndufv2-null*) as a platform for determining the pathogenicity of human mutations in the genes encoding NDUFB10

and NDUFV2, respectively. The human mutations were reconstructed in *Chlamydomonas*, and their effect on complex I activity and assembly were assessed.

2.3 Materials and methods

2.3.1 Strains and culture conditions

Chlamydomonas strains were grown in Tris-acetate-phosphate (TAP), with Hutner's trace elements, 20 mM Tris base and 17 mM acetic acid, or TAP supplemented with arginine (1.9 mM) (TARG), TARG supplemented with 25 μg/ml hygromycin B (TARG + HyB), or 25 μg/ml paromomycin (TARG + Pm) liquid or solid medium at 25°C in continuous light at 50 μmol m^{-2} s^{-1} [1]. In accordance with our laboratory conditions, we define high light conditions as 50 μmol m^{-2} s^{-1} and low light conditions correspond to 0.5 μmol m^{-2} s^{-1}. Solid medium contains 1.5% (w/v) select agar (Invitrogen, #30391049). The background strains used to generate transformants were 3A+ (*mt^+*; *arg7-8*) [CC-5589] and 4C− (*mt^-; arg7-8*) [CC-5590] (Dr. Rochaix, University of Geneva). The strains 141 (*arg9-2; mt^+*), CC-124 (*mt^-*), CC-125 (*mt^+*), or 1' (*mt^+*) [a 137C derivative, provided by Dr. Claire Remacle, University of Liège, Belgium] were used in crosses and/or as experimental controls. Strains *amc5(87D3)* [CC-5591], *dum11* [CC-4098], and *dum18* were used in this study [8, 66]. *Saccharomyces cerevisiae* strain CW04 (*MATα ade2-1 his3-11,15 leu2-3,11 trp1-1 ura3-1* [67]) was utilized for plasmid construction via gap repair described below and grown at 28°C in synthetic dextrose medium containing all

amino acids (SD + AA) prior to plasmid construction. Colonies carrying the recombinant plasmids were selected in synthetic dextrose medium lacking only uracil (SD-ura) [68]. Chemically competent *Escherichia coli* DH5α strains were used for molecular cloning. *E. coli* was grown at 37°C in Luria-Bertani (LB) broth and agar [69].

2.3.2 Insertional mutagenesis and phenotypic screening of complex I mutants

Transformation of 4C- (*mt⁻; arg7-8*) [CC-5590] strain was conducted after autolysine treatment by electroporation [70]. The strain was grown in liquid TARG medium for 2-3 days until exponential growth (3 – 6 x 10^6 cells/mL). For each transformation, 2.5 x 10^7 cells were electroporated at 1.3 kV and 10 μF with 100 ng of hygromycin B resistance cassette (iHyg3) and 20 μg of herring sperm DNA. The iHyg3 cassette consists of the *APHVII* gene with an *RBCS2* intron, under the control of the β2 tubulin-encoding gene (*TUB2*) promoter and the *RBCS2* terminator. This cassette was amplified from the pHyg3 plasmid [71] using the primers: APH7-F and APH7-R (Table S1 of Subrahmanian, *et al*. 2020 [72]). The transformants were selected on TARG+HyB solid medium. Candidate complex I-deficient mutants were identified by scoring for the SID phenotype [8]. Hygromycin B resistant (HyB^R) colonies, which appeared after 10 days of incubation, were transferred into 400 μl of TARG + HyB liquid medium in 96-well plates, each well containing a single colony. The 96-well plate was incubated for five days at 25°C in continuous light at 50 μmol m^{-2} s^{-1}. The liquid cultures were then replica-plated onto two solid TARG media, incubated in the light and dark for five days. Transformants

with a SID phenotype were selected and sub-cloned three times to obtain a single colony prior to further analysis. The *amc* mutant strains isolated in this study and their corresponding *Chlamydomonas* Resource Center reference numbers are as follows: *amc8(1H5)* [CC-5600], *amc9(41D9)* [CC-5601], *amc10(12C)* [CC-5606], *amc11(10G11)* [CC-5608], *amc12(6E9)* [CC-5611], *amc13(4C3)* [CC-5612], and *amc13(16)* [CC-5613].

2.3.3 Ten-fold dilution series and growth curve analysis

The SID phenotype was tested by ten-fold dilution series as follows. One loop of cells grown on solid TARG plates (for 3-5 days) was resuspended in 500 µl of liquid TARG medium. The cell density was measured spectrophotometrically at 750 nm and normalized to an OD750 = 2.0 by dilution. This normalized suspension was used as the starting material (10^0) for making five serial ten-fold dilutions (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}). A volume of 8 µl for each dilution was plated on solid TARG plates. For scoring the SID phenotype, two plates were prepared simultaneously and incubated at 25°C, one in continuous light and another in continuous dark, for at least 7 days. The light-incubated plate served as a control for confirming equal cell density amongst multiple strains.

To determine the generation time, liquid cultures were inoculated with a starting cell density of 10^5 cells mL^{-1} in 50 mL TARG medium. Growth in liquid culture was observed by evaluating cell density at A750. Measurements were taken every 8 hr over a period of 10 days. For each strain, three biological replicates were inoculated in continuous light at 25 µmol m^{-2} s^{-1} or in continuous darkness. Growth rate µ was calculated as 3.3 x

$[\log_{10}N - \log_{10}N_0] / (t_N - t_0)$, where N is the final cell density at time t_N and N_0 is the initial cell density at time t_0. The generation time was calculated as $1/\mu$ [1, 73].

The time periods used for calculating doubling times are as follows: For wild-type (4C-) and [*amc9; NUO5*] strains in the light t_N = 86 hr and t0 = 26 hr, and in the dark t_N = 110 hr and t_0 = 38 hr; for the *amc9* strain in the light t_N = 110 hr and t_0 = 38 hr, and in the dark t_N = 170 hr and t_0 = 86 hr; for wild-type (3A+) in the light t_N = 96 hr and t_0 = 24 hr; for wild-type in the dark t_N = 120 hr and t_0 = 36 hr; for *amc5* in the light t_N = 85 hr and t_0 = 37 hr; for *amc5* in the dark t_N = 123 hr and t_0 = 49 hr; for [*amc5; NUOB10*] in the light t_N = 85 hr and t_0 = 37 hr and for [*amc5; NUOB10*] in the dark t_N = 123 hr and t_0 = 37 hr.

2.3.4 Genetic analysis of the *amc* mutants

Genetic crosses were done as in [1] with some modifications. Vegetative cells were resuspended in TAP liquid medium lacking nitrogen (TAP-N), at 25°C in low light at 0.5-1 μmol m^{-2} s^{-1}, with shaking for a minimum of 5 hr. Equal volumes of the resultant gametes (of opposite mating types) were mixed and incubated in 50 μmol m^{-2} s^{-1} light at 25°C overnight to accomplish mating. In some cases, 10 mM dibutyryl-cAMP (WJ Pharmaceuticals, #100862) and/or 1 mM IBMX (Sigma, #15879-1G) were added to the mixture to stimulate mating. The mixture was plated on TAP-N solid medium (containing 3% (w/v) select agar) and incubated in high light at 25°C for 5 days, to allow the maturation of meiotic zygotes. Meiosis was induced by transferring the zygotes on TARG solid

medium, in high light at 25°C, for at least 12 – 16 hr. The meiotic progeny was obtained through bulk germination or tetrad dissection of zygotes.

In this study, the background strains 141 (*mt*$^{+}$*; arg9-2*), CC-125 (137C+), or 1' (*mt*$^{+}$) were crossed with the original *amc* mutants. Meiotic *amc* progeny derived from these crosses were also used for experimentation. Especially for *amc10* and *amc13*, their meiotic derivatives *amc10(12C)* and *amc13(16)* were used for analyses. For constructing *amc*/+ diploids, the *amc* mutants (*mt*$^{-}$; *amc*; *APHVII; arg7-8*) were crossed with CC-125 and the mating mixture was directly plated on selective medium (TAP + HyB). Individual diploids or meiotic progeny were subcloned to a single colony and their mating type was determined by diagnostic PCR [74].

2.3.5 TAIL-PCR and PCR-based screening of indexed cosmid library

Nucleic acid extraction, diagnostic PCRs, and real-time quantitative PCRs were conducted as described below. TAIL-PCR (Thermal Asymmetric Inter-Laced PCR) was conducted to identify the sequence flanking the iHyg3 cassette (encoding the *APHVII* gene conferring hygromycin B resistance) in the *amc9* mutant as done previously [75] using the partially degenerate primer AD1 [9] (Table S1 Subrahmanian, *et al*. 2020 [72]). The following iHyg3-specific primers, APH7R3, APH7R4, and APH7R5 (Table S1 Subrahmanian, *et al*. 2020 [72]), were used for the primary, secondary, and tertiary TAIL-PCRs, respectively. Similar reactions were conducted using wild-type genomic DNA and purified iHyg3 cassette to identify non-specific amplification of DNA. Cosmids containing

NUO5 and *NUOB10* genomic DNA were identified by PCR [76]. The *NUO5*-containing cosmid (referred to as 9A2) was identified using the primer pairs NUO5 E2L/NUO5 E3R (Table S1). The *NUOB10*-containing cosmid (referred to as cosmid 7D10) was identified using the primer pairs NUOB10E1L/NUOB10E4R (Table S1). The borders of *Chlamydomonas* genomic DNA inserted into these cosmids were sequenced to confirm the presence of genomic region including the gene of interest.

2.3.6 Biolistic transformation

The list of plasmids and recipient strains used for biolistic transformation is provided below and in Tables S2, S3, and S4 of Subrahmanian, *et al*. 2020 [72]. The recipient strains *amc9 (41D9) (mt*$^{-}$; *nuo5::APHVII*; *arg7-8)* [CC-5601] or *amc5 (87D3) (mt*$^{+}$; *nuob10::APHVIII*; *arg7-8)* [CC-5591] were subjected to biolistic transformation using a homemade particle delivery device. The recipient strain was grown in liquid TARG medium for 2–3 days until it reached the exponential phase (3–6 × 10^6 cells/ml). The cells were plated on respective selective medium at 10^8 cells/plate. For each bombardment, DNA was coated on sterile 0.6–0.9 μm tungsten particles (Strem Chemicals, # 93–7437) by using 2 μg of transforming DNA, 16.7 mM spermidine, and 1 M $CaCl_2$. The bombardment was conducted at a helium pressure of 1.725 MPa and vacuum of ~92 kPa. The plate was positioned 10.5 cm away from the particle-containing nozzle. The bombarded plates were first incubated at 0.5 μmol m^{-2} s^{-1} light overnight for recovery and then transferred to continuous light (50 μmol m^{-2} s^{-1}). Transformants containing cosmids

with the *ARG7* marker (9A2 for *amc9* and 7D10 for *amc5*) were selected based on arginine prototrophy. Transformants containing the mutant genes were selected based on their respective antibiotic resistance (Table S4 of Subrahmanian, *et al*. 2020 [72]) and were subsequently screened for the presence of the transgene by diagnostic PCR. The site-directed mutations in each selected transformant were confirmed by sequencing.

2.3.7 Blue-native PAGE (BN-PAGE) and in-gel activity assays

Partially purified membranes were extracted as described above for activity measurement. Complexes were separated by BN-PAGE using 4%–12% (w/v) acrylamide gradient gels [77]. Membranes were partially solubilized as follows. Membrane proteins (500 μg) were pelleted at 18,000 *g* for 20 min at 4°C. The membranes were resuspended in 180 μl of 2% (w/v) sodium *n*-dodecyl-β-D maltoside (DDM; Bioworld, #40430017-3) and solubilized by incubating in DDM in wet ice for 1 hr, followed by addition of 20 μl of 10% (w/v) sodium taurodeoxycholate hydrate (TDC; Sigma, #T-0875). Both DDM and TDC were dissolved in ACA buffer (750 mM aminocaproic acid, 0.5 mM EDTA, 50 mM Bis-Tris, pH 7.0). Partially solubilized membrane proteins (200 μg) were loaded per lane. *In-gel* NADH dehydrogenase (complex I) activity was visualized as purple bands after incubating the gels in 100 mM MOPS-KOH buffer, pH 8, containing 1 mg/ml *p*-nitro blue tetrazolium chloride (NBT; GoldBio, #NBT2.5) and 0.2 mM NADH. Following *in-gel* complex I staining, *in-gel* ATPase (complex V) activity was detected by incubating the gels overnight in the dark, in 50 mM HEPES-KOH pH 8.0 buffer containing 30 mM $CaCl_2$

and 8.2 mM ATP (Fisher Bioreagents, #BP413-25), until a white precipitate was visible. This precipitate revealed the ATPase activity of complex V. Coomassie staining was conducted for loading control.

2.3.8 Immunoblotting analysis

BN-PAGE was completed as described above. For SDS-PAGE, 10 μg of crude membrane proteins was separated by 12.5% acrylamide gel and immunoblotting was performed according to established protocols [78]. The separated proteins were electro-blotted onto PVDF membranes and custom-ordered rabbit polyclonal antibodies specific for *Chlamydomonas* complex I subunits (from Genescript, as described in [8], were used). Membranes with electroblotted proteins separated in BN-PAGE gels were probed overnight at 4°C with 1:3000 diluted α-51 kDa, a polyclonal antibody that detects the soluble arm 51 kDa subunit. Membranes containing proteins resolved by SDS-PAGE immunoblots were probed at room temperature for 3 hr with 1:3000 diluted α-51 kDa, 1: 3000 diluted α-49 kDa, 1:2000 diluted α-TYKY and for 1 hr with 1:12,000 diluted α-cytochrome *f* [79]. For rabbit polyclonal antibodies, an HRP-conjugated anti-rabbit Goat IgG (Biorad, #170-6515) at 1:10,000 dilution was used as the secondary antibody. For detecting the FLAG-tagged proteins, membranes were incubated with monoclonal anti-FLAG antibody (Sigma, #F1804) at 1:5000 dilution overnight at 4°C followed by a HRP-conjugated anti-mouse Goat IgG (Pierce #18584-3) at a dilution of 1:2,500 as the secondary antibody. Intensity of bands in immunoblots were quantified using ImageJ [80].

2.3.9 Complex I activity measurements

Mitochondrial enzymatic activity measurements were conducted as described previously with slight modifications [24, 66, 81]. Cells grown for 2–3 days on solid medium were harvested and resuspended in MOPS KOH extraction buffer (10 mM MOPS-KOH pH 7.4, 0.5 M mannitol, 100 mg/ml BSA, 0.5 mM PMSF). Cells were lysed by sonication using a Branson Sonifier 150 (1/8 inch probe tip), at 12 W output for 2 × 30s. Following lysis, the extract was centrifuged at 480 *g* for 10 min, followed by 3,000 *g* for 5 min. The supernatant was centrifuged at 27,000 *g* for 20 min, and the resulting pellet was the crude membrane extract. Complex I activity was determined as the rate of NADH oxidation, which was measured spectrophotometrically at 340 nm. The substrates used were 100 μM NADH (Amresco, #0384-1G) and 100 μM duroquinone (Aldrich, #D22320-4). Specific activity was calculated using the molar extinction coefficient for NADH at $\varepsilon 340nm = 6.22\ mM^{-1}\ cm^{-1}$ in the absence and presence of 45 μM rotenone (MP Biomedicals, #150154), a complex I-specific inhibitor.

2.3.10 Complex II+III and Complex IV enzymatic assays

Activity assays were conducted on crude membrane extracts that were prepared as described complex I enzymatic activities. Complex II+III activity assay was conducted in the presence of 20.25 mM succinate (Acros organics, #158751000), 1 mM KCN

(FisherScientific, #P223I-100) and 56 μM equine heart cytochrome *c* (Sigma, #2506-500mg). The activity was determined as the rate of cytochrome *c* reduction and measured spectrophotometrically at A550. Complex II+III activity was calculated using molar extinction coefficient for cytochrome *c* at $\Delta\varepsilon 550nm = 19.6\ mM^{-1}\ cm^{-1}$, in the absence and presence of complex III-specific inhibitor myxothiazol (3 μM) (Sigma, #T5580). Complex IV activity assay was conducted in the presence of 1% Triton X-100 and 56 μM reduced cytochrome *c*. Cytochrome *c* was reduced with two times the amount of sodium dithionite and purified with a PD10-desalting column with Sephadex G-25 resin (GE Lifesciences, #17085101) according to the manufacturer's protocol. Complex IV activity was calculated using molar extinction coefficient for cytochrome *c* at $\Delta\varepsilon 550nm = 19.6\ mM^{-1}\ cm^{-1}$, in the absence and presence of complex IV inhibitor KCN (1 mM).

2.3.11 Nucleic acid extraction, diagnostic PCRs and real-time qPCRs

Genomic DNA was extracted from *Chlamydomonas* by the phenol-chloroform method with some modifications [78]. One or two loops of cells, grown for two to three days in continuous light at 50 μmol $m^{-2}\ s^{-1}$ on TARG solid medium, was harvested and resuspended in buffer (10 mM Tris-HCl, pH 8, 10 mM EDTA, 10 mM NaCl, 15% (w/v) glycerol). Cells were lysed by sonication for 5 s at 9 W output. Proteins were degraded by treatment with 100 μg proteinase K (Invitrogen, #25530049) and the extract was incubated at 55 ºC for one hour. RNA was degraded with 50 μg RNase A (Amresco, #0675-250mg).

Nucleic acids were extracted twice by phenol-chloroform and DNA was precipitated by adding 2.2 volumes of ethanol and 0.1 volume of 3 M Na acetate pH 5.5.

The sequences of primers used for diagnostic PCR are shown in Table S1 of Subrahmanian, *et al*. 2020 [72]. For diagnostic PCR analysis, GoTaq Polymerase (Promega, #M3008) was used as recommended by the manufacturer. To enhance the amplification of GC-rich regions, 2.5% (v/v) DMSO was used in each reaction in addition to a denaturation temperature of 98 ºC (instead of 95 ºC). For sequencing analyses, PCR products were gel-purified using the NucleoSpin Gel Extraction Kit (Machery Nagel, #740609.25) as per the manufacturer's instructions. The purified PCR product was then cloned into pGEM-T Easy Vector Systems (Promega, #A1360) and then sequenced with T7 and SP6 primers flanking the cloning site. For real-time quantitative PCR, RNA was prepared as in [82]. RNA was extracted from 2 x 10^8 cells grown in liquid culture. Nucleic acids were extracted twice with equal volume of phenol-chloroform (pH 5.0) and RNA was precipitated after overnight incubation at -20°C with 1/3rd volume of 8 M LiCl. Eight micrograms of RNA were treated with RQ1 RNase-free DNase I (Promega, #M6101). Reverse transcription was achieved with 400 units of M-MLV Reverse transcriptase (Life Technologies, #28025-013) using one microgram of Random Hexamers (Promega, #C1181), following the manufacturer's protocol. Amount of cDNA equivalent to 50 ng or 100 ng of total input RNA was used as template for qPCR using SensiMix (Bioline, #QT-650-05) on a Mastercycler ep gradientS realplex thermocycler (Eppendorf). The sequences of primers used in qPCR are detailed in Table S1. The primer pairs NUO5E2L2 / NUO5E3R and NUOB10E4L / NUOB10E4R were used to amplify *NUO5* and *NUOB10*

mRNA, respectively. The transcript abundance of the housekeeping genes *CBLP* (*Cre13.g599400*) and *EIF1A* (*Cre02.g103550*) were determined using the primers CBLP-F / CBLP-R and EIF1A-Fw / EIF1A-Rev, respectively. Transcript abundance of E3 ubiquitin ligase *UBI* (*Cre03.g159200*) was determined by using primers Ubiupper / Ubilower and TUA2 (Cre04.g216850) using Tub-F / Tub-R [83]. Primer efficiencies for primers binding to *NUO5*, *NUOB10*, *TUA2*, *CBLP*, *UBI*, and *EIF1A*, determined by calibration curves, were 95%, 100%, 100%, 108%, 93%, and 99.8%, respectively. The target transcript was normalized to three reference transcripts by the Livak 2-ΔΔCt method [84]. Three independent biological replicates were performed, each analyzed using two technical replicates. The qPCR reactions were denatured at 98ºC for 15 s, annealed at 60ºC for 20 s and extended at 72ºC for 20 s. Relative fold change was determined by normalizing to the average of the isogenic wild-type strain. Results are represented as percentage of wild-type strain.

2.3.12 Plasmid construction

Plasmids expressing the *NUO5* and *NUOB10* wild-type and mutant genes were generated by recombination-based cloning in yeast. The pRS426-*ble* plasmid, containing three markers allowing for selection in algae, yeast, and bacteria, was used as the vector [85]. It contains the *ble* gene conferring zeocin resistance for selection in *Chlamydomonas*, the *URA3* gene for selection in *S. cerevisiae*, and the *bla* gene conferring ampicillin resistance in *E. coli*. The pRS426-*ble* plasmid was linearized with *Not*I and *Ale*I (pB-NA).

The *NUO5* and the *NUOB10* genes were cloned between the *Not*I and *Ale*I sites in pRS426-ble. The *NUO5* and *NUOB10* genes were expressed under their putative native promoters by including the entire intergenic region upstream of each gene (1220 bp and 1075 bp for *NUO5* and *NUOB10*, respectively). Each gene, including its promoter region, was amplified from purified genomic DNA using Velocity DNA Polymerase (Bioline, #BIO21098) as consecutive overlapping fragments. The sequence corresponding to the C-terminal FLAG-tag and site-directed mutations were introduced by PCR with appropriately designed primers (Table S2 of Subrahmanian, *et al*. 2020 [72]). The overlapping PCR fragments and linearized vector (pB-NA) were assembled via in vivo molecular recombination in yeast (Table S3 of Subrahmanian, *et al*. 2020 [72]). The *S. cerevisiae* strain CW04 [67] was used for recombination. Two hundred nanograms of each linear fragment was introduced into the strain by the one-step transformation method [86, 87] and cells containing the recombinant clones were selected based on uracil prototrophy. Successful recombinants were identified based on diagnostic PCR and restriction digestion, verified by sequencing, and introduced into the respective *Chlamydomonas* strains by biolistics. However, the *ble* selection in *Chlamydomonas* was not successful as it led to the emergence of spontaneous zeocin-resistant colonies in which the transforming DNA was absent. Hence, the *ble* selection marker was substituted with the *APHVII* selection marker (iHyg3) conferring HygB resistance (for the *NUOB10* clones) or *APHVIII* selection marker (iPm) conferring Pm resistance (for the *NUO5* clones). The *ble* marker (1.18 kb) was excised from each plasmid with *Eco*RI and *Not*I restriction enzymes. The iHyg3 and iPm markers were amplified from the plasmids pHyg3 and pSL18 [71, 88],

respectively, using Velocity DNA Polymerase (Bioline, #BIO21098) with primers including a 25 bp overlap with the digested vector (Table S2 of Subrahmanian, *et al*. 2020 [72]). Cloning of the new selection markers were achieved using the In-Fusion HD Cloning Kit (Clontech, #639648) according to the manufacturer's protocol (Table S3 of Subrahmanian, *et al*. 2020 [72]). The recombinant clones finally used for *Chlamydomonas* transformation are: pRS426-iPm-*NUO5*-WT-FLAG and pRS426-iPm-*NUO5*-K230R-FLAG for the *amc9(41D9)* strain; pRS426-iHyg3-*NUOB10*-WT-FLAG, pRS426-iHyg3-*NUOB10*-C79S-FLAG, pRS426-iHyg3-*NUOB10*-C91S-FLAG, and pRS426-iHyg3-*NUOB10*-C79SC91S-FLAG for the *amc5(87D3)* strain (as detailed in Tables S3 and S4 of Subrahmanian, *et al*. 2020 [72]).

2.4 Results

2.4.1 Isolation of novel complex I mutants via forward genetics

To uncover additional *AMC* loci, insertional mutagenesis was conducted using the 4C− wild-type strain (*mt*$^{-}$; *arg7-8*) as the recipient and the iHyg3 cassette, encoding the *APHVII* gene that confers hygromycin B resistance (HyB^{R}), as transforming DNA. The resulting transformants were screened by replica plating for the SID phenotype, a characteristic phenotype of complex I deficiency in *Chlamydomonas* [27]. Among 4,200 insertional mutants, six *amc* mutants (*amc8* to *amc13*) displaying a SID phenotype (Figure 1a) and deficient in rotenone-sensitive NADH: duroquinone oxidoreductase activity

(Figure 1b), were isolated. While the *amc8*, *amc9*, and *amc11* strains had severely decreased complex I activity, the *amc10*, *amc12*, and *amc13* exhibited partial complex I deficiency. To test whether the mutation in the *amc* strains yielded defects in other respiratory enzymes, we measured complex II + III and IV activities (Figure 1c). Apart from *amc12*, all the *amc* mutants exhibited elevated complex II + III activity, a common feature previously observed in several *Chlamydomonas* complex I mutants [8, 24, 27]. None of the *amc* mutants, except *amc12*, displayed a defect in complex IV activity (Figure 1d). We concluded that all *amc* mutants, with the exception of *amc12*, displayed isolated complex I deficiency. The *amc12* mutant was pleiotropic with defects in complexes I, II + III, and IV.

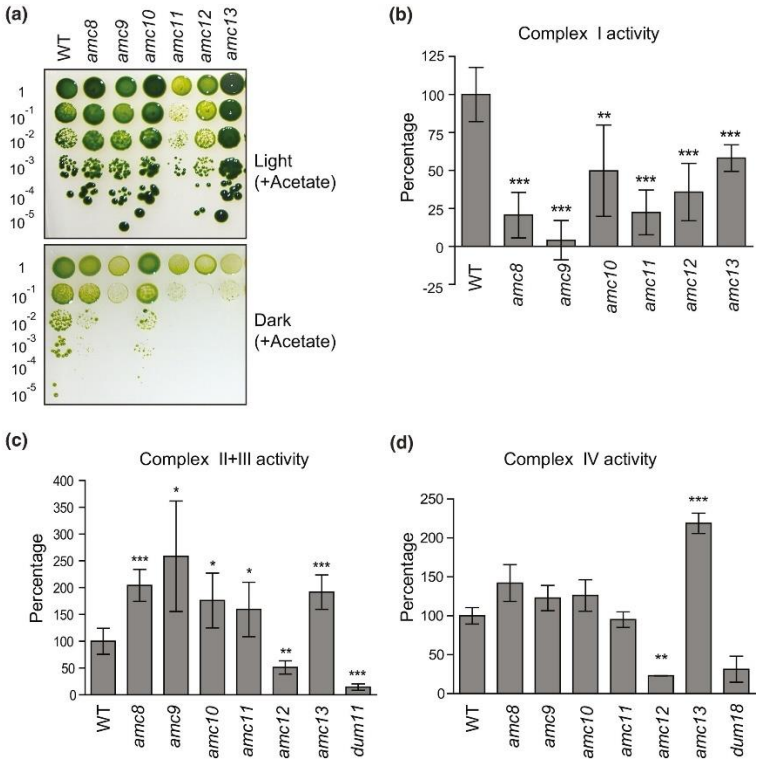


Figure 1 The *amc8* to *amc13* mutants exhibit complex I deficiency.

(a) The growth phenotype of the wild-type (WT, 4C−) and *amc8* to *amc13* mutants was analyzed by 10-fold dilution series. The dilutions were plated on medium containing acetate as a carbon source and incubated in continuous light or in the dark for 20 days. In (b), (c), and (d), the enzymatic activities were conducted on crude membrane extracts and are displayed as percentage of the activity mean of WT, with the error bars indicating percentage of standard deviation of the mean. WT strain used for comparison is 4C− for the *amc* strains and CC-124 for *dum11* and *dum18*. Statistical significance was determined by two-tailed unequal variances t test. * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$. (b) Complex I (rotenone-sensitive NADH: duroquinone oxidoreductase) activity was determined from six independent biological replicates. The average complex I activity of WT was 46.6 ± 8.3 nmol NADH oxidized. min^{-1} mg^{-1} protein. (c) Complex II + III (succinate: cytochrome *c* oxidoreductase) activity was assessed from six independent biological replicates (except *amc8* for which five biological replicates were used). The WT displayed an activity of 18.1 ± 4.3 nmol of cytochrome *c* reduced. min^{-1} mg^{-1} protein. A mutant displaying complex III deficiency (*dum11*) was used as a control. (d) Complex IV activity (cytochrome *c* oxidase) was determined from three independent biological replicates. The WT displayed an activity of 269.6 ± 28.6 nmol of cytochrome *c* oxidized. min^{-1} mg^{-1} protein. A mutant displaying complex IV deficiency (*dum18*) was used as a control. In all the figures, the original mutant strains were used except for *amc10(12C)* and *amc13(16)* (in a, c, and d), which are derivatives of the original *amc10* and *amc13* mutants.

2.4.2 The complex I mutants display defects in complex I assembly

To assess the level of complex I assembly in the newly isolated *amc* mutants, protein complexes from crude membrane extracts were separated via BN-PAGE (blue-native polyacrylamide gel electrophoresis). Mature complex I (~950 kDa) and partially assembled subcomplexes were visualized by *in-gel* staining that reveals NADH dehydrogenase activity as a purple band (Figure 2a). Note that complex I mutants with an assembled soluble arm are capable of *in-gel* NADH oxidation, even if the ubiquinone reductase activity is impaired. Based on the *in-gel* activity, we categorized the *amc s*trains into four groups: (a) no active complex I in the *amc9* mutant, (b) accumulation of a subcomplex displaying NADH dehydrogenase activity in the *amc11* strain, (c) decreased levels of active complex in *amc8 and amc12*, and (d) wild-type levels of *in-gel* NADH dehydrogenase activity in *amc10* and *amc13* strains. BN-PAGE immunoblotting analysis (Figures 2b and 3) showed that no assembled complex I was detected in *amc9*, whereas fully assembled complexes accumulating to a lesser degree than wildtype were observed for *amc8*, *amc10*, *amc12*, and *amc13*. The highly labile subcomplex observed in *amc11* is indicative of a defect in assembling the distal membrane arm of complex I [8, 22]. It has been previously observed that some complex I mutants accumulate reduced levels of complex I subunits as a result of impaired holoenzyme assembly [8, 89]. Hence, the steady-state accumulation of a subset of complex I subunits was examined by SDS-PAGE immunoblotting analysis (Figure 2c). Three subunits from the soluble arm of complex I,

49 kDa (NUO7), 51 kDa (NUO6), and TYKY (NUO8) [8, 36], were chosen for analysis based on the availability of antibodies. Only *amc8, amc11*, and *amc12* accumulated reduced levels of 49 kDa, 51 kDa, and TYKY subunits, whereas *amc10* and *amc13* accumulated these subunits to wild-type levels. The *amc9* mutant displayed decreased levels of the 51 kDa subunit.

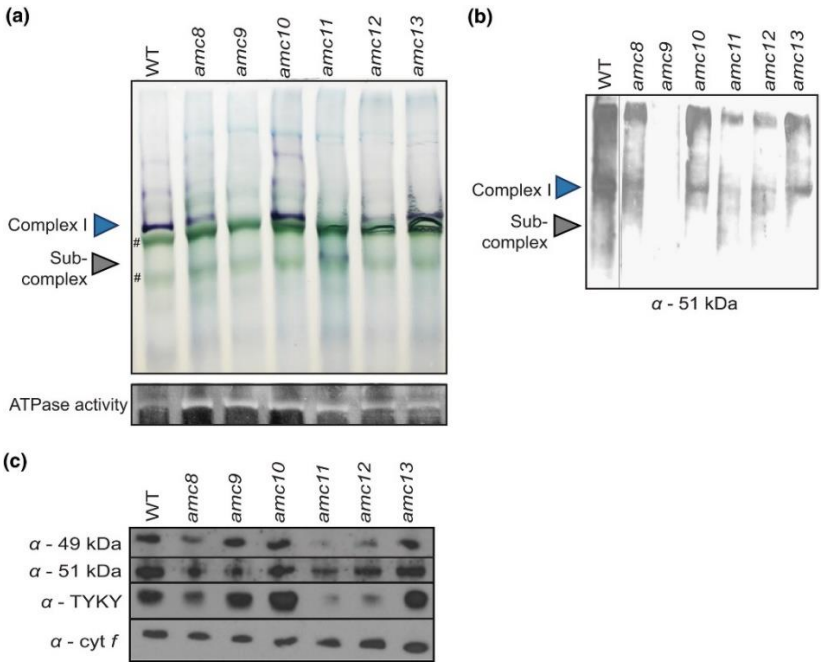


Figure 2 The *amc* mutants display a complex I assembly defect.

(a and b) BN-PAGE (blue-native polyacrylamide gel electrophoresis) was conducted on 200 μg of partially purified membrane fraction. (a, Top panel) In-gel complex I activity was detected by NBT staining. The purple bands indicate in-gel staining of NADH dehydrogenase activity in mature (950 kDa) and partially assembled subcomplexes of complex I. In some cases, multiple purple bands larger than 950 kDa are detected, which could be due to partially solubilized membranes or might correspond to complex I in association with other complexes [22]. The symbol (#) indicates the photosynthetic complexes present in the crude membrane extract, migrating at sizes similar to that of the complex I holoenzyme and the subcomplex. They are marked only in two lanes for ease of reference. (a, Bottom panel) In-gel ATPase staining to detect complex V was conducted to verify quality of crude membrane preparation and loading. The reduced ATPase staining in WT compared to the *amc* mutants was not systematically observed. (b) Immunoblotting was conducted, using α-51 kDa antibody, on complexes separated by BN-PAGE to detect the assembled soluble arm of complex I. This image is a composite of two gels run on the same day as indicated with a black vertical line: with the WT lane from one gel and the lanes corresponding to the *amc* mutants from another. (c) SDS-PAGE immunoblotting was conducted on 10 μg of partially purified membranes using polyclonal antibodies to detect soluble arm complex I subunits: α-49 kDa, α-51 kDa, α-TYKY. α-cyt *f* was used to confirm equal loading. In (a) and (b), WT is the 4C− strain and the *amc8* to *amc13* strains are the original mutants except the *amc10(12C)* strain, which is a meiotic progeny derived from the original *amc10* strain.

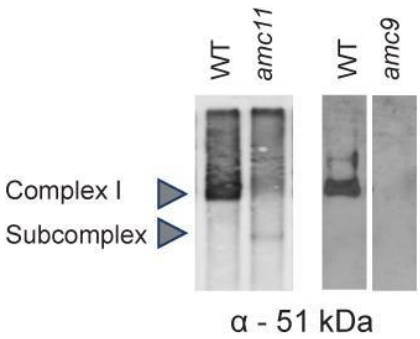


Figure 3 Additional Blue-Native PAGE immunoblotting analyses.

Blue-Native PAGE followed by immunoblotting was conducted on crude membrane extracts using a polyclonal antibody to detect the 51 kDa subunit of complex I. The white vertical line between the lanes of WT and *amc9* denotes the assembly of lanes from the same blot.

2.4.3 The *amc9, amc11,* and *amc12* mutations are linked to the insertional marker

Although the complex I mutants were generated by insertional mutagenesis, previous genetic analyses of *amc* mutants showed that the insertional cassette is not always linked to complex I deficiency [8]. The unlinked mutations could be due to the insertion of extracellular genomic DNA uptaken during electroporation [90], insertion of cleaved and non-functional pieces of the cassette, or insertion of the random herring sperm DNA used as part of the electroporation protocol [70]. Hence, genetic analyses were conducted to determine the nature of the *amc* mutations. Analysis of the heterozygous diploid progeny (*amc*/+) showed the diploids were restored for growth in the dark, indicating that all *amc* mutations were recessive for the complex I-deficient phenotype (Figure 4). Meiotic progeny of *amc* × wild-type crosses were tested to determine whether the *amc* mutations were monogenic and the insertional cassette co-segregated with the complex I-deficient phenotype (Table 1). The analyses indicated that the complex I deficiency in the *amc8*, *amc9*, *amc10*, *amc11*, and *amc13* strains exhibited monogenic inheritance (Table 1). In addition, all the HyB^R meiotic progeny, derived from genetic crosses of wild-type, with *amc9*, *amc11*, or *amc12*, displayed a SID phenotype, indicating that the complex I deficiency in these mutants is tightly linked to the insertional cassette. On the other hand, for *amc8*, *amc10*, *and amc13*, only a fraction of the HyB^R recombinant meiotic progeny displayed the SID phenotype, indicating that the *AMC* locus responsible for the complex I-deficient phenotype was segregating away from the antibiotic resistance insertional cassette. Tetrad analysis also confirmed that the complex I deficiency was not linked to the

insertional cassette in the *amc10* and *amc13* mutants. We conclude that the recessive complex I deficiency in the *amc* mutants was linked to the insertional cassette only for *amc9, amc11,* and *amc12.* In this study, we show further characterization of the *amc9* mutant.

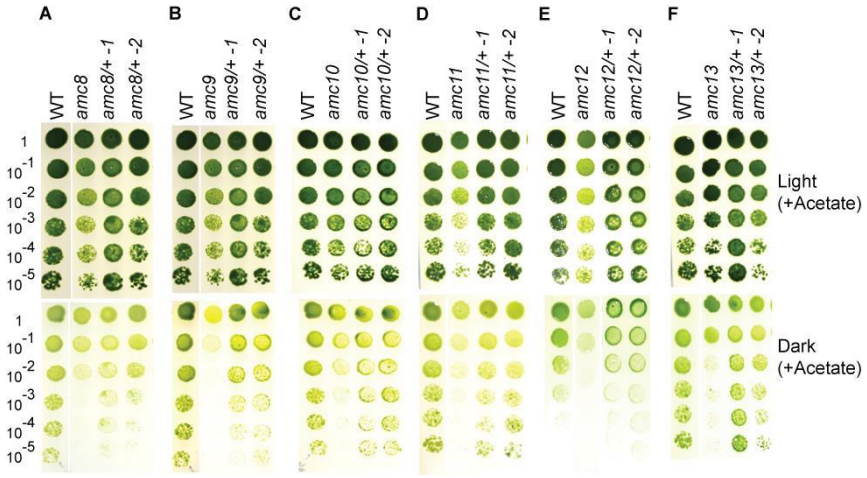


Figure 4 The *amc8* to *amc13* mutations are recessive.

To test if the *amc* mutations causing the complex I-deficient phenotype are recessive or dominant with respect to the wild-type allele, heterozygous diploids (*amc*/+) were constructed by crossing *amc* mutants with a wild-type strain as described in the materials and methods. Two independent diploids from each cross were tested for growth in the dark by ten-fold dilution series. The dilution series were plated on acetate-containing medium and incubated in the light or dark for 10 days. The heterozygous *amc*/+ diploids were restored for growth in the dark, thereby indicating that all the *amc* mutations are recessive. White vertical lines indicate pictures of strains grown from the same plate and assembled for the display in the figure. The growth of the following diploids is depicted in panels A-F wherein WT (4C-) and the respective *amc* mutants, used for constructing the diploids, were used as controls: A, *amc8*/+; B, *amc9*/+; C, *amc10*/+; D, *amc11*/+; E, *amc12*/+; F, *amc13*/+.

Table 1 Phenotypic and genetic analysis of the *amc* mutants.

Strain	CI activity (%)	Fully assembled complex	Subcomplex	Genetic analysis	Recombinant meiotic progeny		Linkage to cassette	Monogenic
					Total	SID		
WT	100	++++	–	–	–	–	–	–
amc8	21	++	–	Bulk	230	50	No	Yes
amc9	4	–	–	Bulk, Tetrad	50	50	Yes	Yes
amc10	50	+++	–	Bulk, Tetrad	100	48	No	Yes
amc11	22	–	++	Bulk	100	100	Yes	Yes
amc12	36	++	–	Bulk	85	85	Yes	N.D.
amc13	58	+++	–	Bulk, Tetrad	112	51	No	Yes

Complex I-specific activity for the amc mutants was determined by measuring rotenone-sensitive NADH: duroquinone oxidoreductase activity and is represented as percentage of WT activity (WT, 4C− at 46.6 ± 8.3 nmol NADH oxidized. min^{-1} mg^{-1} protein, set to 100%). The detection of fully assembled complex and the subcomplex was determined from BN-PAGE (blue-native polyacrylamide gel electrophoresis) in-gel activity and immunoblotting. ++++, +++, ++, +, − indicate relative levels of detected complex. To test whether the *amc* mutations are monogenic, genetic analysis of the meiotic progeny of *amc* x wild-type crosses was performed by analyzing the 2:2 segregation of complex I phenotype in tetrads. In cases where tetrad analysis was not successful, bulk germination of zygotes was conducted and the resulting meiotic progeny were scored for complex I-deficient phenotype. In the case of *amc9*, *amc10*, and *amc13*, each of the tetrads that were tested showed a 2:2 segregation of the SID (slow growth in the dark) and wild-type heterotrophic growth phenotype, confirming monogenic inheritance of the complex I-deficient trait (Figure 5). In the case of the *amc12* mutant, the monogenic inheritance of the SID phenotype could not be determined (N.D.) via tetrad analysis due to poor germination of the zygotes.

2.4.4 The *amc9* mutation maps to the *NUO5* gene encoding the 24 kDa subunit of complex I

In the case of *amc9*, analyses of seven tetrads and 50 HyB^R recombinant meiotic progeny obtained from bulk germination of meiotic zygotes showed that the antibiotic resistance always segregated with the SID phenotype (Figure 5). The tight linkage between the insertional marker and the SID phenotype in the *amc9* mutant suggests that the disruption of a gene controlling complex I, by the insertional cassette, could be responsible for the complex I deficiency. To identify the disrupted gene in the *amc9* mutant, we sought to recover the genomic sequence flanking the insertional cassette via TAIL-PCR [75]. The full-length insertional cassette was mapped to exon 2 of the *NUO5* gene, which encodes the 24 kDa subunit (NUO5) of the soluble arm of complex I (Figure 15a,b) [28]. Real-time RT-qPCR (Figure 15c) showed that the *amc9* mutant lacked the full-length *NUO5* transcript. These results confirmed the insertion of a full-sized iHyg3 cassette into exon 2 of the *NUO5* gene in the *amc9* mutant.

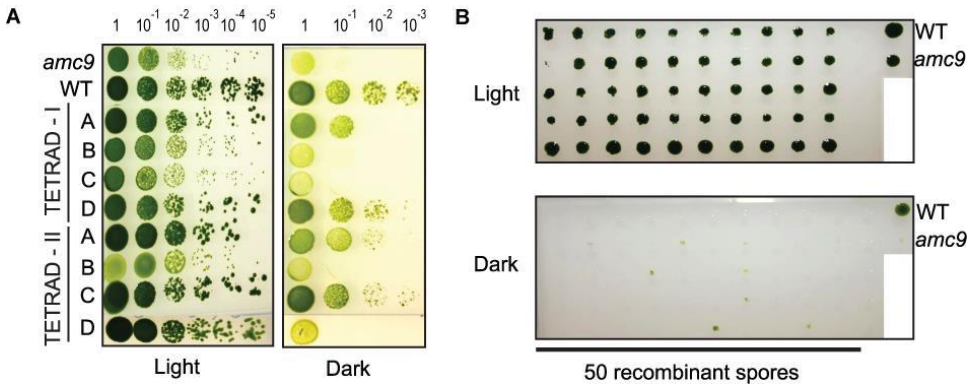


Figure 5 The *amc9* mutation is linked to the insertional cassette.

Meiotic zygotes were obtained by crossing the complex I proficient strain 141 (*mt*$^{+}$*; arg9-2*) with the complex I-deficient *amc9(3-)* strain (*mt*$^{-}$*; nuo5::APHVII*). The meiotic zygotes, obtained from the same genetic cross, were used for both tetrad dissection (A) and bulk germination (B). (A) The growth phenotype of two out of seven tetrads dissected from the above-mentioned cross is shown here by ten-fold dilution series. The dilutions were plated on acetate-containing medium and incubated in the light and in the dark for 15 days. The WT strain and the *amc9* strain shown here are the original parental strains. (B) Fifty recombinant hygromycin B resistant progeny, were analyzed from bulk germination of meiotic zygotes. The progeny were replica-plated on TARG solid medium in the light or dark and the growth phenotype was scored after 15 days of incubation. The WT and *amc9* strains shown here are 4C- and *amc9(41D9)*, respectively. Out of 50 antibiotic-resistant meiotic progeny that were tested, all of them displayed a SID phenotype confirming that the *amc9* mutation is tightly linked to the insertional cassette.

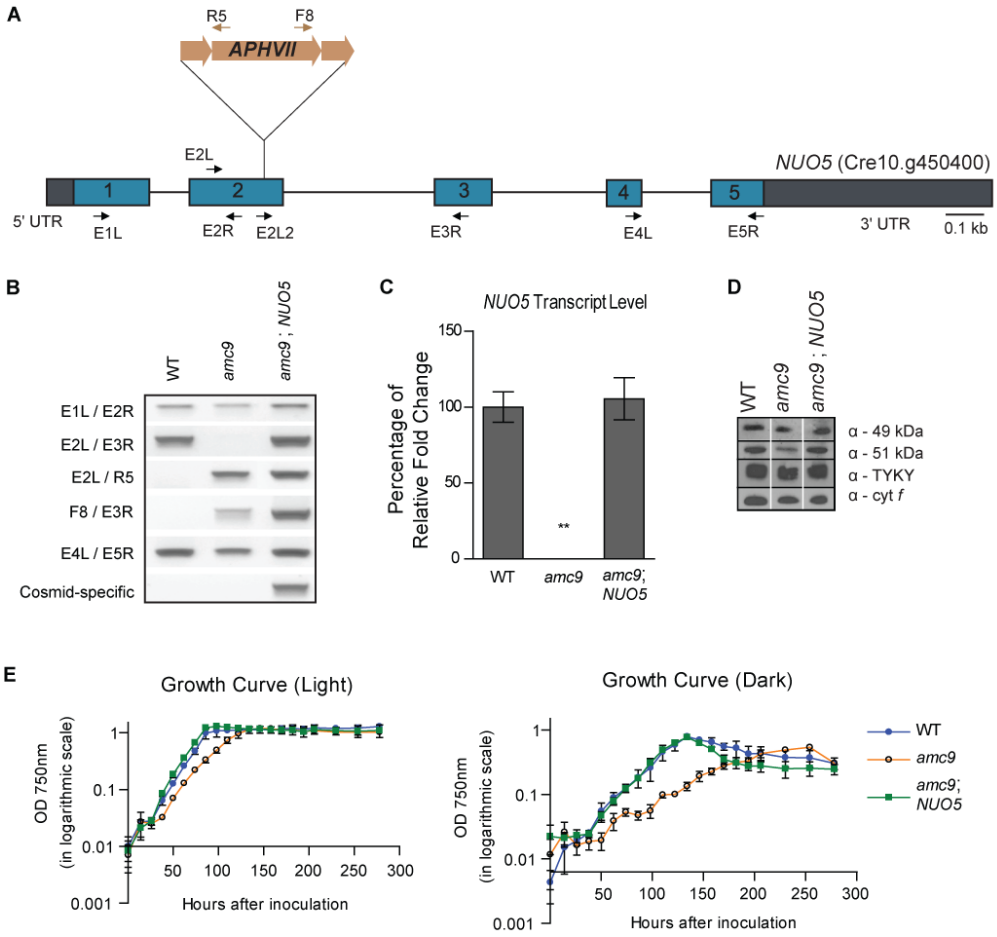


Figure 6 The *NUO5* gene encoding the 24 kDa complex I subunit is disrupted in the *amc9* strain.

(A) A diagram of the *NUO5* gene with the approximate position of the insertional cassette in the *amc9* mutant is depicted here. The brown arrows indicate primers APH7R5 and APH7F8 that bind to the *APHVII* gene in the iHyg3 cassette. The black arrows represent *NUO5*-specific primers NUO5E1L, NUO5E2R, NUO5E2L, NUO5E2L2, NUO5E3R, NUO5E4L, and NUO5E5R (Table S1 of Subrahmanian, *et al*. 2020 [72]). (B) Diagnostic PCR of the *NUO5* gene in wild-type (WT, 4C-), *amc9(41D9)*, and one *amc9* strain transformed with a cosmid containing the *NUO5* gene [*amc9; NUO5*] is shown here. The positions of the primers used for diagnostic PCR of the *NUO5* gene are represented in (A). The location of the insertion was confirmed by amplification of the genomic region spanning the insertion site using primer pairs NUO5E2L / NUO5E3R. A band of expected size (715 bp) was amplified in the wild-type and not in the *amc9* mutant. The insertion of the full-length iHyg3 cassette was confirmed by amplifying across the *NUO5*/iHyg3 junctions in the *amc9* mutant. Primer pairs amplifying from the 5′-end (NUO5E2L / APH7-R5) and from the 3′-end (APH7-F8 / NUO5E3R) of the iHyg3 insertional cassette yielded products of expected size, 524 bp and 955 bp, respectively, in *amc9* but not in the wild-

type. Diagnostic PCRs conducted upstream and downstream of the insertion site yielded amplicons of the expected size, indicating the absence of other major rearrangements in the *NUO5* gene of the *amc9* strain. Diagnostic PCR of the [*amc9; NUO5*] transformant revealed the presence of the wild-type *NUO5* gene, in addition to the endogenous disrupted *NUO5* gene. The primers (PH3 and 9A2 PH3, Table S1) were used for determining the presence of the cosmid (last row), harboring the wild-type *NUO5* gene. (C) Real-time quantitative PCR was used to assess the quantity of *NUO5* mRNA relative to the transcripts of three reference genes *TUA2*, *CBLP* and *UBI*. The average was obtained from three biological replicates, each including two technical replicates. The error bars represent standard deviation of the mean. The results are represented as percentage of fold change relative to WT (WT set to 100). Statistically significant difference with respect to the wild-type was determined by two-tailed unequal variances t-test. ** indicates $p < 0.01$. (D) SDS-PAGE immunoblotting was conducted on 10 μg of partially purified membranes using polyclonal antibodies to detect complex I soluble arm subunits: α- 49 kDa, α- 51 kDa, α-TYKY. α– cyt *f* was used to confirm equal loading. Vertical white lines indicate the assembly of different lanes from the same immunoblot. (E) The growth of WT (4C-), *amc9(41D9)* and [*amc9; NUO5*] was recorded by measuring optical density at 750 nm, in the light (22 μmol m^{-2} s^{-1}) or in the dark, over a period of 10 days. The average of three biological replicates is reported here, with error bars indicating standard deviation of the mean.

To test whether the insertional mutation in the *NUO5* gene is responsible for the complex I defect, we transformed the *amc9* mutant with a cosmid containing the wild-type copy of *NUO5* (referred to as [*amc9*; *NUO5*]) and assessed the recovery of complex I function. The *NUO5* transcript levels were restored upon complementation with the wild-type *NUO5* gene (Figure 6c). Rescue of the growth phenotype in the [*amc9*; *NUO5*] strain was measured by assessing growth on solid medium and in liquid culture (Figures 7a,b and 6e). The generation time in the dark for wild-type and *amc9* was 17.5 hrs and 52 hrs, respectively, whereas the [*amc9*; *NUO5*] strain displayed a wild-type level of growth in the dark as evidenced by a generation time of 16.8 hr (Figure 7b). Further biochemical analyses of [*amc9*; *NUO5*] revealed wild-type levels of NADH: duroquinone oxidoreductase activity (Figure 7d), complex I subunits abundance (Figure 6d), and complex I assembly (Figure 7e,f). From these results, we conclude that the *AMC9* locus corresponds to the *NUO5* gene encoding the 24 kDa complex I subunit (referred to as NUO5 in *Chlamydomonas* and NDUFV2 in humans [28, 48].

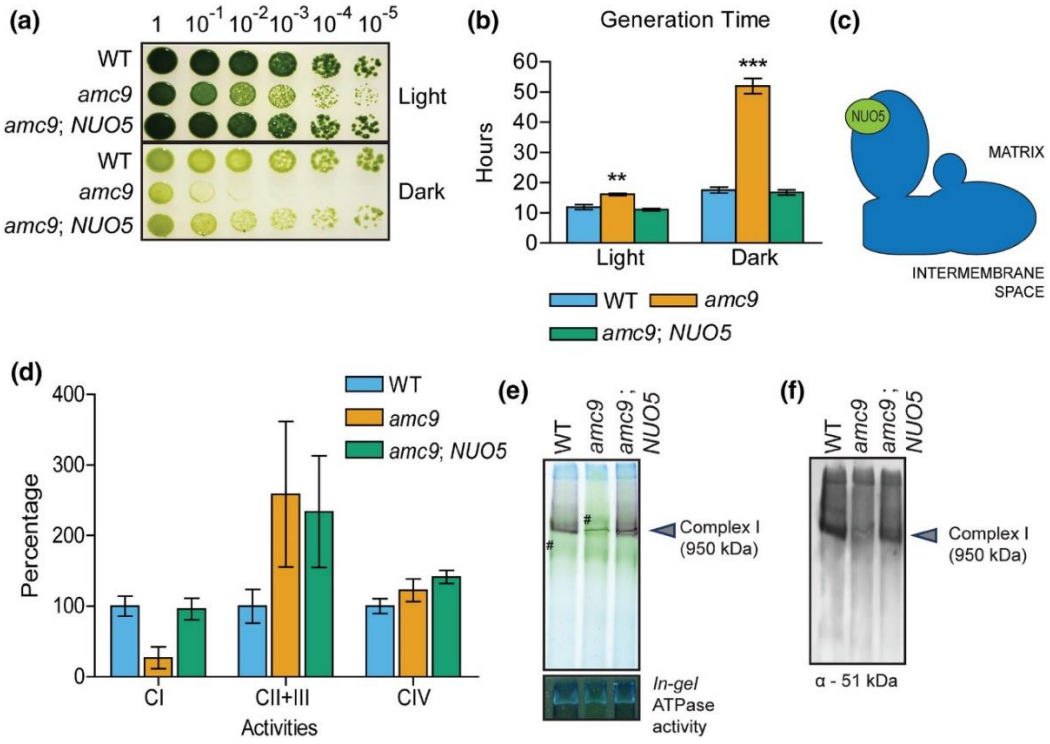


Figure 7 The *amc9* mutant, complemented by the *NUO5* gene, is restored for complex I activity and assembly.

(a) Restoration of the growth phenotype in [*amc9; NUO5*] was tested by 10-fold dilution series plated on acetate-containing medium and incubated in the light for seven days and in the dark for 16 days. (b) The average generation time for each strain calculated from growth curves in Figure 15e is indicated here. The error bars represent standard deviation of the mean. Statistically significant difference with respect to the WT was determined by two-tailed unequal variances t test. ** indicates $p < 0.01$, and *** indicates $p < 0.001$. (c) The approximate location of the NUO5 subunit in the matrix arm of complex I is indicated in a diagrammatic representation. (d) Complex I (CI), complex II + III (CII + III), and complex IV (CIV) activities were determined on partially purified membranes. The activities are represented as percentage of WT calculated from an average, with the error bars indicating standard deviation of the mean. The averages for CI, CII + III, and CIV activities were determined from three, six, and three biological replicates, respectively. For the reference strain, WT (4C−), average CI activity was 76.4 ± 19.9 nmol NADH oxidized. min^{-1} mg^{-1} protein, average CII + III activity was 18.1 ± 4.3 nmol cytochrome *c* reduced. min^{-1} mg^{-1} protein, and average CIV activity was 269.6 ± 28.6 nmol cytochrome *c* oxidized min^{-1} mg^{-1} protein. The *amc9* mutant displays a significant reduction in complex I activity with respect to WT, as determined by two-tailed unequal variances t test with a p value = 0.03. The [*amc9; NUO5*] strain is rescued for complex I activity. While there is no significant difference between activities measured for the WT and [*amc9; NUO5*] strains, there is a significant difference between *amc9* and [*amc9; NUO5*] with a p = 0.019. (e)

BN-PAGE (blue-native polyacrylamide gel electrophoresis) was conducted on 200 μg of partially purified membrane fraction. In-gel complex I activity was detected by NBT staining. The symbol (#) indicates the chlorophyll-containing complexes present in the crude membrane extract. In-gel ATPase activity was detected with $CaCl_2$/ATP staining. (f) BN-PAGE followed by immunoblotting was conducted on 200 μg of partially purified membrane fraction using polyclonal antibody to detect the 51 kDa subunit of the soluble arm of complex I.

2.4.5 The NDUFV2 K209R variant does not affect complex I activity or assembly in *Chlamydomonas*

NDUFV2/NUO5 is the 24 kDa soluble subunit localized to the matrix arm of the holoenzyme (Figure 7c) and is one of the core subunits harboring a 2Fe-2S (N1a) cluster which is coordinated by four cysteines [91] (Figure 8). As a highly conserved protein, human NDUFV2 displays 51% identity with the *Chlamydomonas* NUO5 ortholog (Figure 8) and is a known marker for complex I disorders [48]. While the association of complex I deficiency with Parkinson's disease (PD) is a well-established phenomenon, the exact molecular mechanisms defining how specific complex I-related mutations cause pathogenesis have remained unclear [52]. One particular lysine-to-arginine (K209R) variant in NDUFV2 was detected in one out of 33 familial probands and one out of 238 sporadic PD cases [92]. However, complex I enzymatic activity was not assessed in these patients to determine whether this mutation causes a complex I deficiency that may contribute to the development of PD.

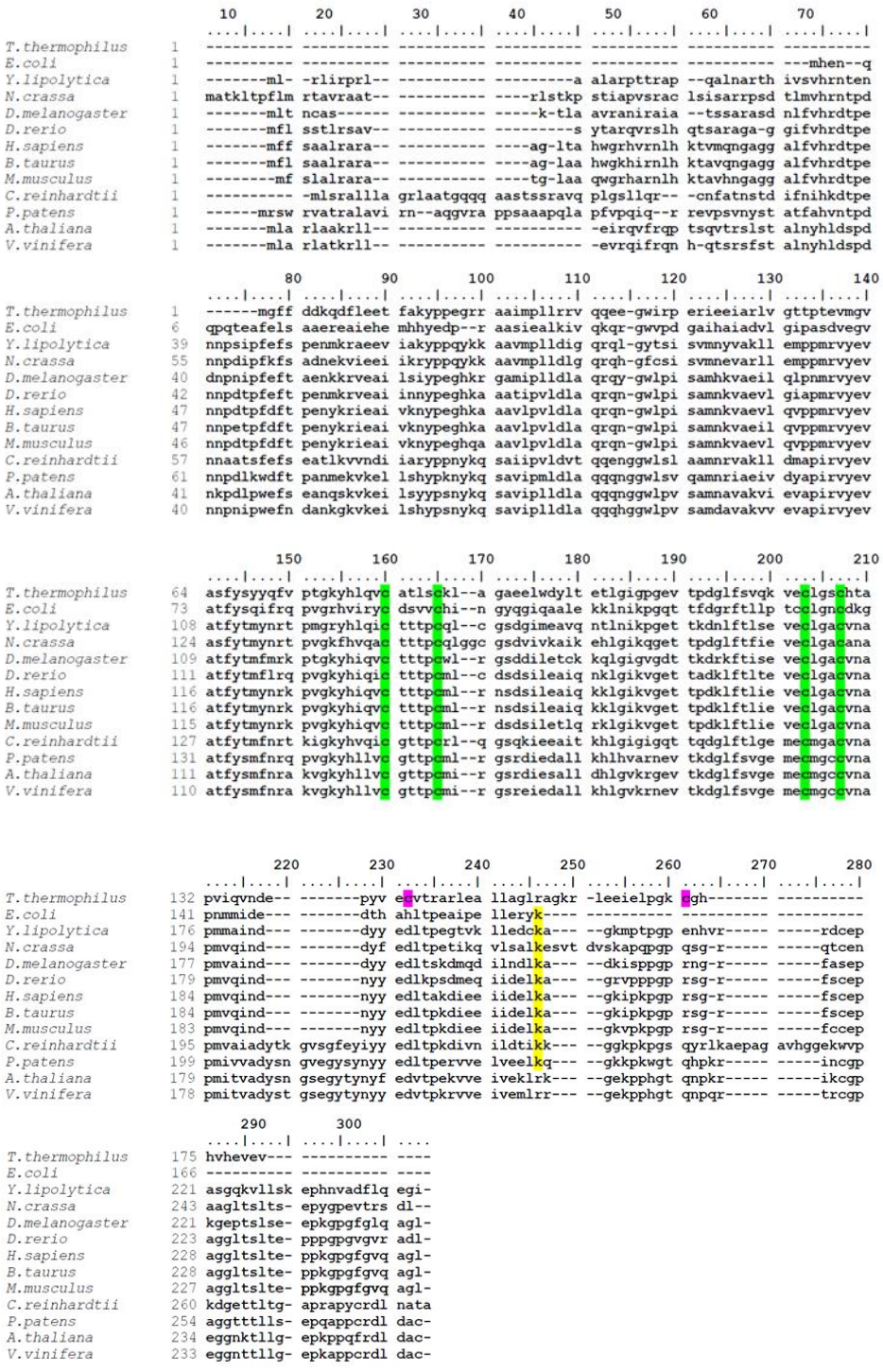


Figure 8 Alignment of NUO5 / NDUFV2 / 24 kDa subunit orthologs.

NUO5 / NDUFV2 / 24 kDa subunit proteins were aligned using Clustal Omega (BLOSUM 62 scoring matrix) and Bioedit [93]. The conserved cysteines that coordinate the 2Fe-2S (N1a) cluster in NUO5 are highlighted in green. In E. coli, the N1a cluster has a high midpoint redox potential implying a possible role in electron transfer through a flavosemiquinone intermediate [94]. However, in mammalian complex I this cluster exhibits a low midpoint redox potential implying inability for electron relay and hence NDUFV2 is alternatively proposed to function in the structural stability of the enzyme [91, 94]. The two cysteines that provide increased stability by forming a disulfide in *T. thermophilus* ortholog are highlighted in magenta [95]. The conserved lysine residue that is mutated to arginine in the Parkinson's disease patients [92] is highlighted in yellow. NCBI accession numbers are as follows: *T. thermophilus* (AAA97942), *E. coli* (EDV69206), *Y. lipolytica* (XP_502254), *N. crassa* (XP_961535), *D. melanogaster* (NP_573228), *D. rerio* (NP_957041), *H. sapiens* (NP_066552), *B. taurus* (NP_776990), *M. musculus* (NP_082664), *C. reinhardtii* (XP_001698508), *P. patens* (XP_001759238), *A. thaliana* (NP_567244), *V. vinifera* (XP_002281655).

In some cases, it has been shown that lysine-to-arginine substitutions may affect protein folding, and in others, it has been proposed that lysine-to-arginine substitutions increase stability through putative salt bridges and hydrogen bond formations [96]. To investigate the biochemical effect of the NDUFV2-K209R variant on complex I activity and assembly, the corresponding mutation was reconstructed in the gene encoding the *Chlamydomonas* NUO5 ortholog and the variant was expressed in the *nuo5-null* mutant strain *amc9*. This lysine residue is well-conserved in eukaryotic species and occurs at position 230 of *Chlamydomonas* NUO5 (yellow highlight, Figure 8). The sequence encoding either the wild-type or the K230R subunit was introduced into a construct containing the *NUO5* genomic DNA, fused to a sequence encoding a C-terminal FLAG-tag, under the control of the *NUO5* native promoter. Transformants were generated in the *amc9* strain via biolistics, and those accumulating the FLAG-tagged NUO5 protein were selected for further analyses (Figure 9).

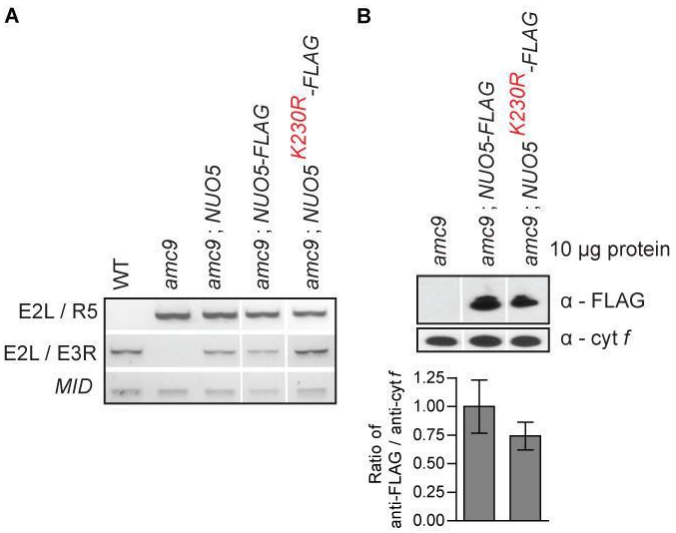


Figure 9 The FLAG-tagged variants of NUO5 are expressed in the amc9 mutant.

The *amc9* mutant was transformed with two constructs containing *NUO5* genomic DNA: i) wild-type *NUO5* sequence (NUO5-FLAG), or ii) mutant *NUO5* sequence encoding the K230R substitution (NUO5K230R-FLAG). The wild-type (WT, 4C-), *amc9(41D9)* mutant and [*amc9; NUO5*] strains were used as controls. White vertical thin lines indicate lanes that have been assembled together from the same gel. (A) Diagnostic PCR of the *NUO5* gene in the *amc9* transformants, at the site of insertion (NUO5 E2L / NUO5 E3R), resulted in amplification of wild-type sequence indicating the presence of transforming DNA. All transformants also contained the *amc9* insertional mutation (NUO5 E2L / APH7R5). The sequence and position of the primers is available in Supplemental Table S1 of Subrahmanian, *et al*. 2020 [72] and Figure 5. Amplification of the *MID* gene was used as control. (B) Immunoblotting was conducted on whole cell proteins separated by SDS-PAGE using monoclonal antibodies to detect FLAG tag (α- FLAG). Polyclonal antibody α– cyt *f* was used to test for equal loading. The proteins were quantified with ImageJ software from three independent biological replicates. No significant difference in NUO5-FLAG accumulation was observed for the wild-type NUO5-FLAG and the K230R NUO5-FLAG variant.

The potential impact of the lysine-to-arginine substitution on growth was assessed (Figure 10a). As observed previously, the *amc9* mutant strain exhibited a SID phenotype whereas complementation with the wild-type *NUO5* gene or recombinant *NUO5* gene with a C-terminal FLAG-tag restored growth in the dark to wild-type levels. Interestingly, transformants expressing the K230R NUO5- FLAG variant also exhibited wild-type growth in the dark, indicating that this substitution in NUO5 does not affect respiratory growth (Figure 10a). Accordingly, NADH: duroquinone oxidoreductase activity was also restored by the K230R NUO5 variant (Figure 10b). To test the level of complex I assembly in the NUO5 lysine-to-arginine variant, complexes were separated by BN-PAGE and subjected to *in-gel* activity assays and immunoblotting (Figure 10c). The NUO5 subunit is part of the matrix arm responsible for NADH dehydrogenase activity. As a result, no mature complex I or subcomplexes were detected by *in-gel* activity or immunoblotting in the *amc9* mutant. Complementation with the wild-type *NUO5-FLAG* gene restored complex I assembly. In addition, immunoblotting with the α-51 kDa or α-FLAG antibodies showed that the NUO5-FLAG K230R variant was successfully incorporated into the holoenzyme. This was in agreement with detection of fully assembled complex I via *in-gel* activity assay showing that the NADH dehydrogenase activity of the soluble arm was restored. From these results, we conclude that the K230R substitution in NUO5, a candidate mutation for PD, does not affect complex I activity or assembly in *Chlamydomonas.*

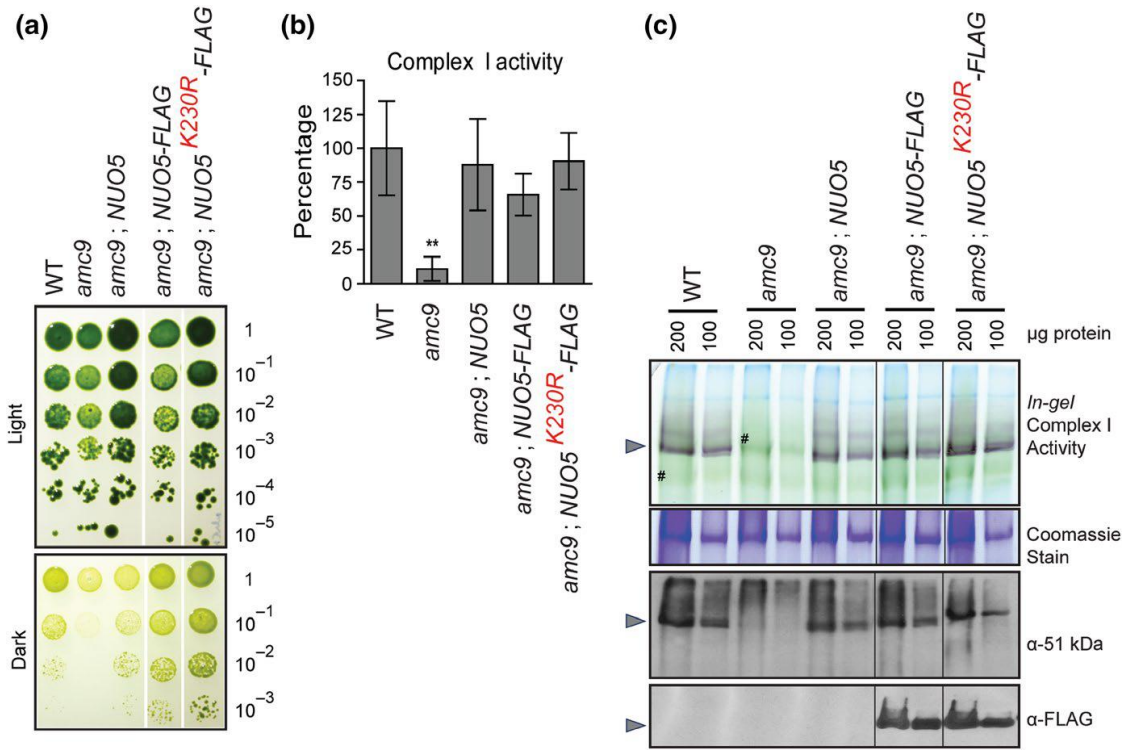


Figure 10 The lysine-to-arginine substitution in NUO5 does not affect complex I activity and assembly.

(a) The growth phenotype of the WT and the *amc9* transformants was analyzed by 10-fold dilution series plated on acetate-containing medium and incubated in the light or in the dark for 14 days. White vertical lines indicate strains tested on the same plate and assembled together for this figure. (b) Complex I activity was determined with partially purified membranes from four biological replicates and represented as a percentage of WT average with the error bars indicating standard deviation of the mean. For WT (4C−), average complex I activity was 71 ± 24.7 nmol NADH oxidized. min^{-1} mg^{-1} protein. The *amc9* mutant displayed a significant reduction in complex I activity with respect to WT, as determined by two-tailed unequal variances *t* test with a *p* value = 0.005. The [*amc9*; *NUO5-FLAG*] and the [*amc9*; *NUO5*^{K230R}*-FLAG*] transformants were restored for complex I activity. (c) BN-PAGE (blue-native polyacrylamide gel electrophoresis) was conducted on 200 µg and 100 µg of partially purified membrane fraction. *In-gel* complex I activity was detected by NBT staining. The gray arrowheads indicate fully assembled holoenzyme. The green bands (indicated with the symbol #) correlate to the co-purified photosynthetic complexes in the membrane fractions. Coomassie staining following BN-PAGE was used to test for equal loading. Protein(s) migrating at a size unrelated to complex I is shown here for this purpose. BN-PAGE followed by immunoblotting was conducted using a polyclonal antibody to detect the 51 kDa subunit of the soluble arm of complex I and a monoclonal α-FLAG antibody to detect NUO5-FLAG. Black vertical thin lines indicate lanes assembled together from the same gel/blot. The NUO5^{K230R}-FLAG is incorporated into the fully

assembled complex I and allows wild-type levels of NADH dehydrogenase activity as detected by *in-gel* complex I activity assay.

2.4.6 The *AMC5* locus corresponds to the NUOB10-encoding gene

NUOB10/NDUFB10/PDSW is an accessory subunit that is localized to the distal end of the membrane arm, although NUOB10 is hydrophilic and does not contain any predicted transmembrane helices [35] (Figures 11e and 12). It is predicted to face the intermembrane space (IMS) and presumably associated with the membrane through interactions with neighboring membrane subunits [97]. The *Chlamydomonas* NUOB10 displays 13.5% identity with the human ortholog NDUFB10.

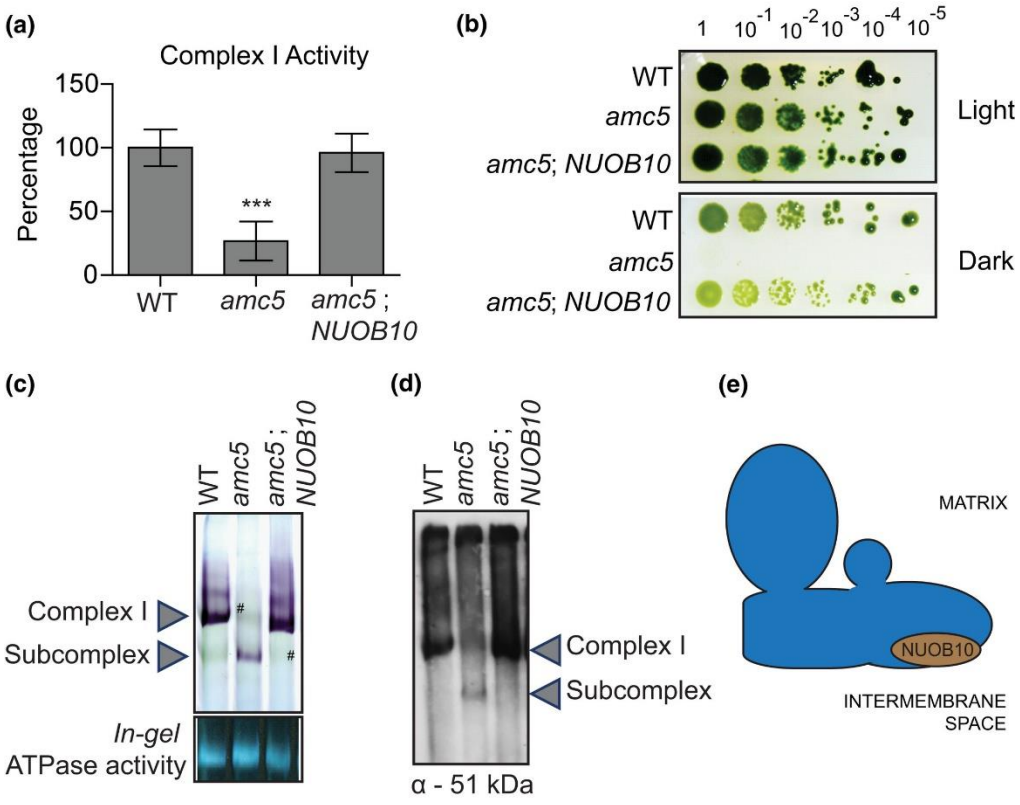


Figure 11 The amc5 mutant phenotypes are rescued by the NUOB10 gene.

The *amc5* strain was transformed with a *NUOB10*-containing cosmid by biolistics. The phenotypic rescue of one representative transformant [*amc5*; *NUOB10*] is shown here. The WT and *amc5* strains shown here are 3A+ and *amc5 (87D3)*, respectively. (a) Complex I (rotenone-sensitive NADH: duroquinone oxidoreductase) activity was determined with partially purified membranes. The activities are represented as a percentage of the average of three biological replicates with the error bars indicating percentage of standard deviation of the mean. The average complex I activity of WT is 54.0 ± 7.7 nmol NADH oxidized. min^{-1} mg^{-1} protein. The activities for the WT and *amc5* strains are significantly different according to the two-tailed unequal variances *t* test with a $p = .000116$. The *amc5* and [*amc5*; *NUOB10*] strain have significantly different activities with a $p = .000199$. (b) Restoration of the growth phenotype in [*amc5*; *NUOB10*] was tested by 10-fold dilution series plated on acetate-containing medium and incubated in the light for seven days and in the dark for 16 days. (c and d) BN-PAGE was conducted on 200 μg of partially purified membranes. (c) *In-gel* complex I activity was detected by NBT staining, and *in-gel* ATPase activity was detected with $CaCl_2$/ATP staining. Crude membrane extracts, also containing the photosynthetic membranes, were used for this analysis. The symbol (#) indicates examples of photosynthetic complexes migrating closely with complex I and subcomplex that can be clearly observed in their absence. (d) Immunoblotting was conducted using a polyclonal antibody to detect the 51 kDa subunit of complex I. (e) A diagrammatic

representation of the L-shaped mitochondrial complex I, with the approximate location of the NUOB10/NDUFB10 subunit in the distal membrane arm facing the IMS [97].

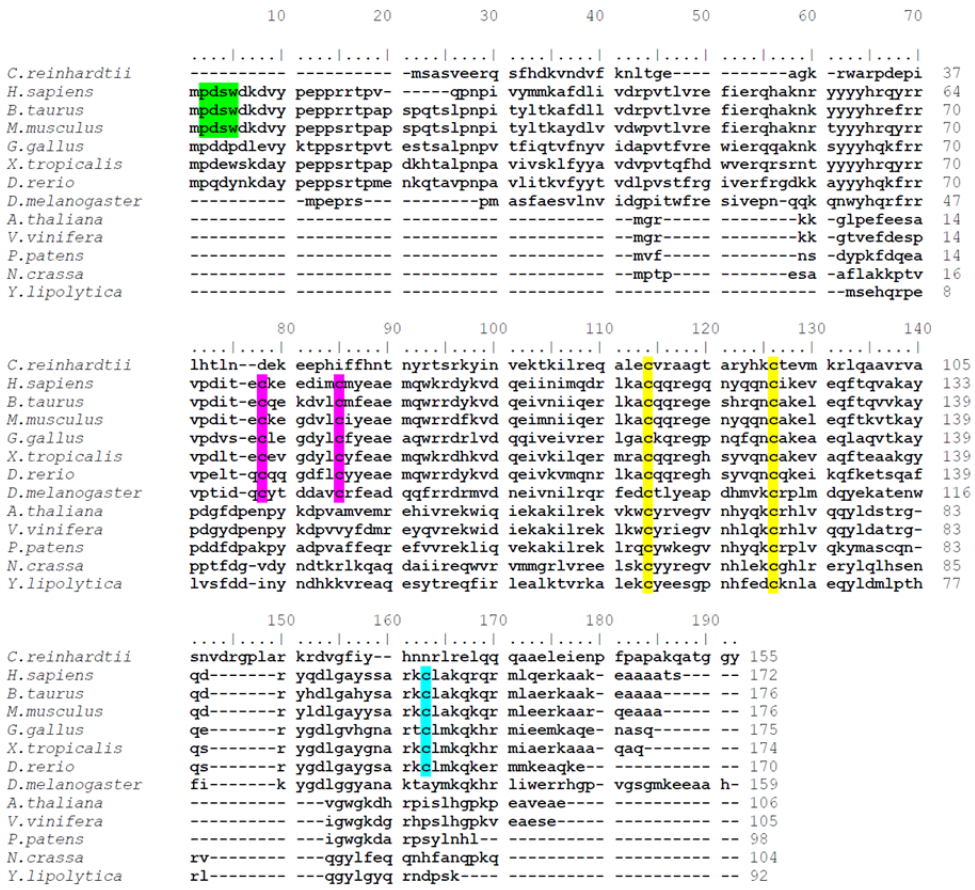


Figure 12 Alignment of NUOB10 / NDUFB10 / PDSW subunits orthologs.

NUOB10/NDUFB10/PDSW subunit proteins were aligned using Clustal Omega (BLOSUM 62 scoring matrix) and Bioedit [93, 98]. The conserved PDSW sequence at the N-terminus is highlighted in green. The two cysteines of the C-$(X)_{11}$-C motif are highlighted in yellow, the C-$(X)_6$-C motif in pink and the fifth cysteine in the human ortholog are highlighted in blue. NCBI accession numbers are as follows: *D. melanogaster* (NP_651972), *G. gallus* (XP_414844), *H. sapiens* (O96000), *B. taurus* (Q02373), *M. musculus* (Q9DCS9), *X. tropicalis* (NP_001017250), *D. rerio* (Q6PC16), *C. reinhardtii* (XP_001694041), *P. patens* (XP_001770976), *A. thaliana* (NP_566608), *V. vinifera* (XP_002276810), *N. crassa* (CAA48768), *Y. lipolytica* (XP_002142991).

The *amc5* mutant harbors the insertional cassette in intron 3 of the *NUOB10* gene [8] accompanied by a deletion of the *NUOB10* genomic sequence downstream of the insertion site (Figure 13a,b). The *amc5* mutant displayed decreased rotenone-sensitive NADH: duroquinone oxidoreductase activity (Figure 11a) and exhibited the characteristic *sid* phenotype of complex I-deficient mutants in both liquid and solid medium (Figures 11b and 13d,e) with an average generation time of 69 hr in the dark, compared to 27 hr for the wildtype strain. Real-time RT-qPCR confirmed the loss of the wild-type *NUOB10* mRNA in the *amc5* mutant (Figure 13c). The *amc5* mutant displayed an accumulation of a subcomplex, migrating at a size similar to the ~700 kDa subcomplex previously observed in *Chlamydomonas* mitochondrial mutants defective for the distal membrane arm assembly of complex I [22, 38] (Figure 11c,d).

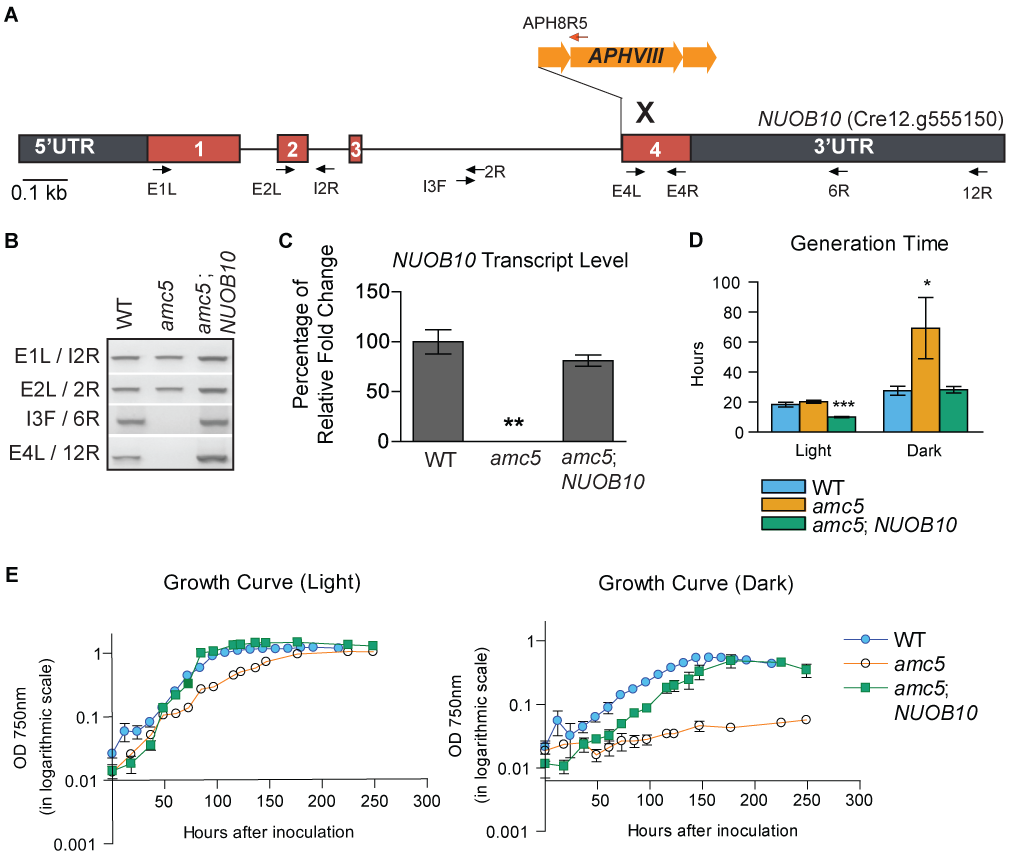


Figure 13 The wild-type NUOB10 gene restores heterotrophic growth to the amc5 mutant.

(A) A diagram of the *NUOB10* gene with the position of the insertional cassette in the *amc5* mutant is depicted here. The gray and red rectangles represent UTRs and the coding sequence of *NUOB10*, respectively. The introns are denoted by thin black lines. The *X* mark indicates deletion of the *NUOB10* genomic region downstream of the insertion site. The black arrows indicate the primer binding sites for the *NUOB10*-specific primers and the orange primer is specific to the insertional cassette (iPm). (B) Diagnostic PCR of the *NUOB10* gene in the *amc5* strain reveals the presence of the molecular lesion as previously described in [8]. The [*amc5*; *NUOB10*] strain contains an intact copy of the *NUOB10* gene as expected. (C) Real - time RT-qPCR was used to assess the quantity of *NUOB10* mRNA relative to three reference genes *TUA2*, *CBLP* and *EIFA*. The average was obtained from three biological replicates, each including two technical replicates. The error bars represent standard deviation of the mean. The results are represented as percentage of fold change relative to WT (WT set to 100). Statistical significance was determined with respect to wild-type by two-tailed unequal variance *t*-test. ** indicates *p*-value < 0.01. All primer sequences described in (A – C) are detailed in Supplementary Table S1 of Subrahmanian, *et al*. 2020 [72]. (D and E) The growth of the wild-type (WT, $3A^{+}$), *amc5* and [*amc5*; *NUOB10*] strains was recorded by measuring optical density at 750 nm, in the light or in

the dark, over a period of 10 days. The average of three biological replicates is reported in (E), with error bars indicating standard deviation of the mean. The average generation time for each strain calculated from growth curves is depicted in (D). The error bars represent standard deviation of the mean. Statistical significance was determined by two-tailed equal variances *t*-test and * indicates *p*-value < 0.05, *** indicates *p*-value < 0.001.

To test whether the mutation in *NUOB10* is indeed responsible for the complex I defect, the *amc5* mutant was transformed with a cosmid containing the *NUOB10* gene (Figures 11 and 13). Molecular analyses of the [*amc5*; *NUOB10*] transformant revealed the presence of the wildtype *NUOB10* gene and restoration of relative *NUOB10* transcript levels. The [*amc5*; *NUOB10*] strain also exhibited restoration of growth in the dark, complex I activity, and assembly. From these results, we conclude that the *AMC5* locus corresponds to the *NUOB10* gene and the NUOB10 subunit is necessary for complex I membrane arm assembly.

2.4.7 The NUOB10 C-$(X)_{11}$-C motif is important for complex I activity and assembly

To date, only one patient has been reported with mutations in *NDUFB10*. The patient, born to non-symptomatic parents, exhibited fetal cardiomyopathy and fatal infantile lactic acidosis, and died at 27 hr after birth [46]. Exome sequencing identified compound heterozygous sequence variation in the *NDUFB10* gene: (a) a paternally inherited nonsense mutation resulting in a premature stop codon, and (b) a maternally inherited missense mutation resulting in a cysteine-to-serine (C107S) substitution. This cysteine is part of a highly conserved C-$(X)_{11}$-C motif (yellow highlight, Figure 12), whose function in complex I activity and assembly has not been elucidated.

To understand the effect of the C107S substitution in NDUFB10 on complex I holoenzyme, we sought to reconstruct the corresponding mutation in the *Chlamydomonas nuob10*-null *amc5* mutant. The C107 residue in human NDUFB10 corresponds to the first

cysteine of the C-$(X)_{11}$-C motif, at position 79 in the *Chlamydomonas* NUOB10 ortholog. To further gain insight into the role of the C-$(X)_{11}$-C motif in complex I activity, a cysteine-to-serine substitution at the second cysteine of the motif (C91S) and the double substitution (C79S-C91S) were also tested. For this purpose, the *NUOB10* genomic sequence (corresponding to wild-type, C79S, C91S, C79S-C91S variants) was fused to a sequence encoding a C-terminal FLAG-tag and expressed under the control of its native promoter. The *NUOB10* constructs were introduced into the *amc5/nuob10*-null mutant strain by biolistics. Transformants carrying the transgene were identified and chosen for further analyses (Figure 14a). To test the accumulation of NUOB10-FLAG in the transformants, immunoblotting was conducted. The NUOB10-FLAG variants were detected at the expected size of ~17 kDa with the α-FLAG antibody (Figure 14b).

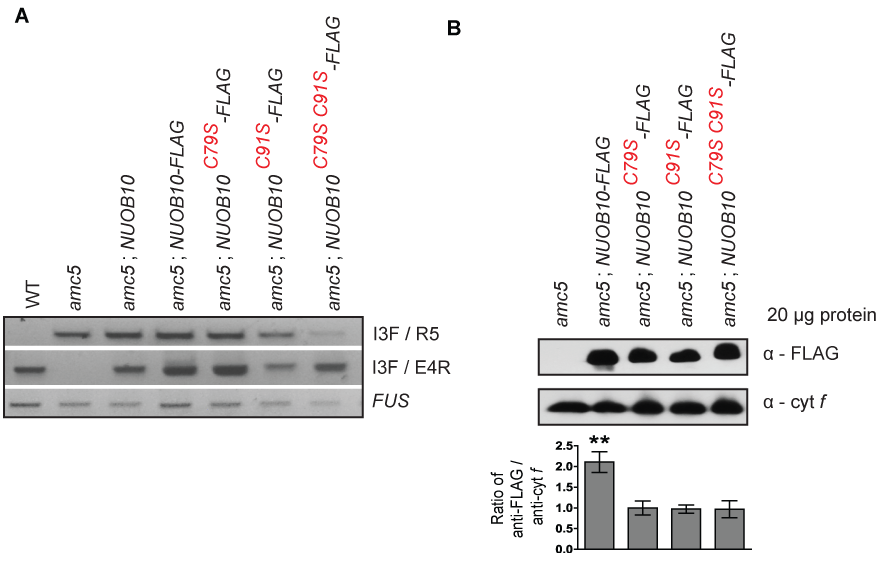


Figure 14 The FLAG-tagged variants of NUOB10 are produced in the amc5 mutant.

The *amc5* mutant was transformed with four constructs containing the *NUOB10* genomic DNA: i) wild-type *NUOB10* sequence (*NUOB10* or *NUOB10-FLAG*), and mutant *NUOB10* sequences encoding the variants with ii) the C79S substitution (*NUOB10*^{C79S}-*FLAG*), iii) the C91S substitution (*NUOB10*^{C91S}-*FLAG*), or iv) the C79S and C91S double substitutions (*NUOB10*^{C79SC91S}-*FLAG*). The wild-type (WT, 3A^{+}), *amc5* mutant and [*amc5*; *NUOB10*] strains were used as controls. (A) Diagnostic PCR of the *NUOB10* gene in the *amc5* transformants, at the site of insertion (NUOB10 I3F / NUOB10 E4R), shows amplification of wild-type sequence indicating the presence of transforming DNA. All transformants also contain the *amc5* insertional mutation (NUOB10 I3F/ APH8R5). The position of the primers is available in Figure 13a. Amplification of the *FUS* gene was used as control. (B) *(top):* SDS-PAGE immunoblotting was conducted on whole cell protein extract using monoclonal α- FLAG antibody to detect FLAG-tagged NUOB10. Polyclonal antibody to detect α– cyt *f* was used to test for equal loading. *(below):*The immunoblots were quantified by ImageJ software from three independent biological replicates [80]. The transformant [*amc5*; *NUOB10-FLAG*] accumulates significantly more NUOB10-FLAG compared to other transformants with a $p < 0.01$ (calculated from three biological replicates using two-tailed unequal variances *t*-test).

The effect of the NUOB10 cysteine substitutions on respiratory growth was tested (Figure 15a). While the *amc5* recipient strain displayed a SID phenotype, transformants expressing the wild-type *NUOB10-FLAG* had restored growth in the dark similar to those with the *NUOB10*-containing cosmid. On the other hand, the single and double cysteine-to-serine variants displayed only partial restoration of growth in the dark. These observations indicate that while expression of the *NUOB10* variants can partially compensate for loss of NUOB10, manipulation of the C-$(X)_{11}$-C motif restricts respiratory growth.

To test the importance of the C-$(X)_{11}$-C motif for complex I activity and assembly, rotenone-sensitive NADH: duroquinone oxidoreductase activity and BN-PAGE *in-gel* activity were assessed (Figure 15b,c). The *amc5* strain transformed with wild-type *NUOB10-FLAG* showed rescue of complex I activity to ~65% of wild-type levels (Figure 15b). Although the restoration of complex I activity for the [*amc5*; *NUOB10-FLAG*] transformant was not as high as with the [*amc5*; *NUOB10*] control strain, BN-PAGE *in-gel* activity and immunoblotting revealed wild-type levels of mature complex I at ~950 kDa, showing that the C-terminal FLAG-tag does not significantly impair complex I assembly (Figure 15c). On the other hand, the cysteine-to-serine single and double substitutions in NUOB10 yielded severe complex I deficiency, indicating that the C-$(X)_{11}$-C motif is crucial for complex I activity (Figure 15b). Interestingly, the accumulation of the 700 kDa subcomplex due to the loss of NUOB10 in the *amc5* mutant was attenuated in the presence of the NUOB10 variants (Figure 15c). Mature complex I (~950 kDa), absent in the *amc5* mutant, was detected in the single and double cysteine-to-serine variants via

in-gel activity and immunoblotting with α-51 kDa. Immunoblotting with α-FLAG (detecting NUOB10) further revealed that the NUOB10 variants were incorporated into the mature complex. The restoration of complex I assembly, in spite of the cysteine-to-serine substitutions, could account for the slight improvement in the complex I activity levels of the cysteine-to-serine variants (~16% of wild-type) compared to the *amc5* mutant (~8% of wild-type) and partial rescue of the respiratory growth phenotype (Figure 15a,b). However, the NUOB10 variants failed to accumulate wild-type levels of mature complex I, as evidenced by immunoblotting (Figure 15b). The level of complex I assembly is similar for the single or double mutant variants, an indication that mutation of either or both cysteines of the C-$(X)_{11}$-C motif elicits the same impact on complex I activity and assembly. Additionally, the three distinct NUOB10-FLAG cysteine- to-serine variants accumulated to only 50% of wild-type (Figure 14b), indicating that the cysteines may also be required for the stability of NUOB10. From these observations, we concluded that the cysteines within the highly conserved C-$(X)_{11}$-C motif of NUOB10 play an important role in complex I assembly and activity.

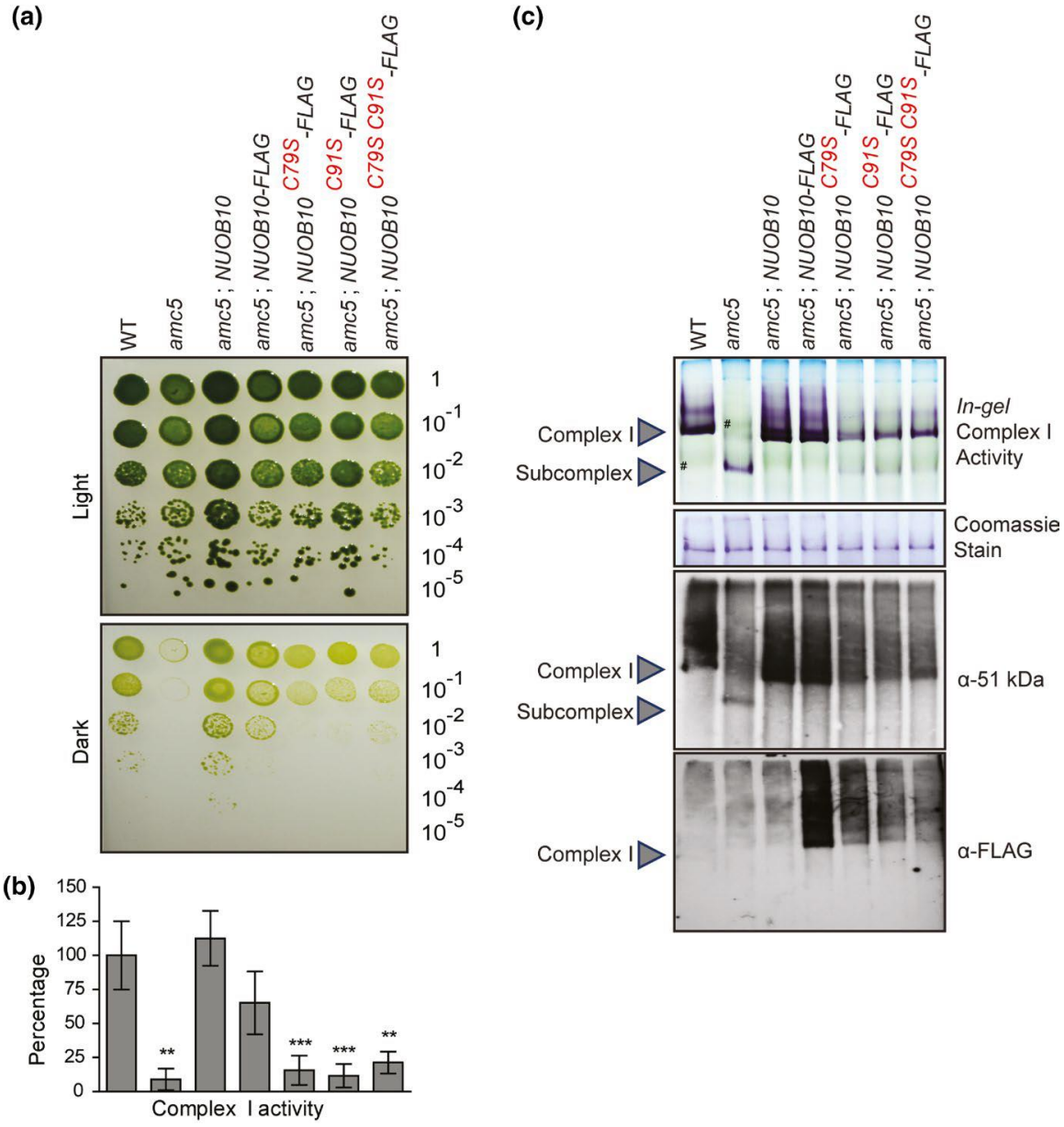


Figure 15 The cysteine-to-serine substitutions in NUOB10 decrease complex I activity and assembly.

The *amc5* mutant was transformed with four constructs containing the *NUOB10* genomic DNA: (i) wild-type *NUOB10* sequence (*NUOB10* or *NUOB10-FLAG*), and mutant *NUOB10* sequences encoding the variants with (ii) the C79S substitution (*NUOB10*^{C79S}-*FLAG*), (iii) the C91S substitution (*NUOB10*^{C91S}-*FLAG*), or (iv) the C79S and C91S double substitutions (*NUOB10*^{C79S,C91S}-*FLAG*). The wildtype (WT, 3A+), *amc5* mutant, and [*amc5*; *NUOB10*] strains were used as controls. (a) The growth phenotype of the wild-type and *amc5* transformants was analyzed by 10-fold dilution series. The dilutions were plated

on medium containing acetate as a carbon source and incubated in the light or in the dark for 14 days. (b) Complex I activities, conducted with partially purified membranes, are represented as percentage of WT calculated from the average of five biological replicates with the error bars indicating standard deviation of the mean. For WT, average complex I activity was 48.3 ± 12.1 nmol NADH oxidized. min^{-1} mg^{-1} protein. The *amc5* transformants producing the C79S and C79S, C91S variants display significantly reduced complex I activities compared to WT as determined by two-tailed unequal variances *t* test. ** indicates $p < 0.01$, and *** indicates $p < 0.001$. The graph is aligned to match with the labels in (a). (c) BN-PAGE was conducted on 200 μg of partially purified membrane proteins. *In-gel* complex I activity was detected by NBT staining. The symbol (#) indicates the photosynthetic complexes present in the crude membrane extract. Although present in all the lanes, they are marked only in two lanes for ease of reference. Coomassie staining serves as loading control, and protein(s) migrating at a size different from complex I is shown here. Immunoblotting was conducted using a polyclonal antibody to detect the 51 kDa subunit of the soluble arm of complex I and a monoclonal α-FLAG antibody to detect the NUOB10-FLAG.

2.5 Discussion

Mitochondrial complex I, the first and largest enzyme of the mitochondrial ETC, is a proton-pumping NADH: ubiquinone biogenesis factors. Considering the number of proteins required for complex I biogenesis [28], it is clear our forward genetic screen is far from saturated. Screening for the SID phenotype appears to have only a 0.02% frequency of obtaining a *bona fide* complex I mutant. The use of a larger insertional mutant library, similar to the CLiP library generated by the Jonikas group [90], could yield additional novel *AMC* loci. Unfortunately, in our experience, the conditions used for generating the CLiP mutants are more conducive for isolating photosynthetic-deficient mutants and less so for respiratory-deficient mutants. The Remacle group has devised a new method of screening for respiratory mutants that is based on the concerted contribution of the photosynthetic and respiratory systems to cellular ATP production [25, 99]. They used the *pgrl1* mutant, defective for photosystem I cyclic electron transfer that is involved in ATP production in chloroplasts, as the background for generating respiratory mutants. Respiratory deficiency in the *pgrl1* background displayed an additional phenotype defined by decreased photosystem II efficiency, which was used as the basis to screen for complex I-deficient nuclear mutants [25, 99]. Again, only three out of 3,059 transformants (0.09%) were identified as true complex I mutants from this screen. Future development of a screen to positively select or enrich for complex I mutants after mutagenesis might increase the success rate, yielding a larger number of nuclear mutants deficient for complex I.

Chlamydomonas has been previously used as a successful tool for studying human mitochondrial mutations. The L158P substitution in the mitochondrially encoded ND4 subunit, observed in one patient with chronic progressive external ophthalmoplegia [100], was shown to affect complex I activity but not assembly when reconstructed in *Chlamydomonas* [62]. In the second part of this study, we exploited the high degree of conservation *of Chlamydomonas* complex I nuclear-encoded subunits with their human counterparts [36, 61] and demonstrated the efficacy of utilizing the newly uncovered *Chlamydomonas* nuclear mutants as a tool for defining the consequence of potentially pathogenic human mutations on complex I assembly and activity.

Both NUO5 (NDUFV2 in human) and NUOB10 (NDUFB10 in human) are highly conserved complex I subunits that have been identified as critical markers in human mitochondrial disorders [46, 92, 101-103]. NDUFV2 has been implicated in Alzheimer's disease, bipolar disorder, Parkinson's disease, and other pathologies. In this study, we tested a provisional mutation causing a lysine-to-arginine substitution in NDUFV2 and showed that it does not affect complex I activity or assembly. Indeed, comparative analyses of NDUFV2 orthologs revealed that an arginine residue is present at this position in *Thermus thermophilus*, *Arabidopsis*, and *Vitis vinifera* (Figure 8). As both lysine and arginine are positively charged amino acids, making the substitution of a conservative nature, it is likely this particular substitution is well-tolerated and does not elicit a change in complex I activity in *Chlamydomonas*. While complex I is a highly conserved enzyme in all eukaryotes, it is expected that some differences do exist between organisms regarding the point of entry of subunits during the assembly process [104]. Hence, we cannot rule out

that the lack of an effect on complex I due to the lysine-to-arginine substitution is *Chlamydomonas*-specific. Three other SNPs in the *NDUFV2* gene resulting in amino acid substitutions of uncertain significance have been recently listed in the ClinVar database: P139T, M185V, and D190G. The P139T and M185V are especially interesting substitutions as these residues are very close to the cysteines involved in binding the Fe-S cluster (Figure 8). A future line of investigation could be to test the importance of these residues with our *nuo5*-null mutant.

The first patient reported with an isolated complex I disorder due to a mutation in the human nuclear *NDUFB10* gene presented an early-onset phenotype, characterized by prenatal cardiomyopathy along with metabolic acidosis and failure to thrive [46]. In concert with a compound heterozygous nonsense mutation, a missense mutation characterized by a cysteine-to-serine substitution of the first cysteine in the C-$(X)_{11}$-C motif resulted in decreased levels of complex I activity and increased accumulation of assembly intermediates in skeletal muscle, heart, and liver tissues [46].

Recently, NDUFB10 was identified as an interacting partner of CHCHD4, a disulfide bond-forming enzyme, via affinity purification in both denaturing and native conditions, implying that NDUFB10 could be an in vivo target of CHCHD4 [105]. The CHCHD4-ALR import machinery (Mia40-Erv1 in yeast), also known as the mitochondrial IMS assembly system (MIA), functions by interacting with the cysteine residues of the substrate proteins and driving their import from the outer mitochondrial membrane into the IMS by coupling translocation with disulfide bond formation [106, 107]. Canonical CHCHD4 substrates contain twin C-$(X)_3$-C or C-$(X)_9$-C motifs, whose cysteines form

intramolecular disulfide bonds spanning the two motifs, enabling the formation of an anti-parallel helix–turn–helix structure [107]. The complex I subunits NDUFS5, NDUFB7, and NDUFA8 are canonical substrates of the CHCHD4-MIA system containing twin C-$(X)_9$-C motifs [108]. Interestingly, the human NDUFB10 protein contains five cysteines in non-canonical motifs: a C-$(X)_6$-C motif, the C-$(X)_{11}$-C motif, and a fifth single cysteine (Figure 12). Pulse-chase experiments showed that cysteine thiols in NDUFB10 are no longer free after import into isolated human mitochondria and therefore presumed to be disulfide-linked [46]. It is expected that two disulfide bonds are formed in NDUFB10 upon import [46]. Although the identity of the disulfide bond-forming cysteines remains to be biochemically ascertained, single-particle electron cryo-microscopy of fungal, murine, bovine, and ovine complexes I model a disulfide bond between the cysteines within each motif of the NDUFB10 subunit [109-112]. As NDUFB10 does not have a canonical mitochondrial targeting sequence that is cleaved upon import [35] and NDUFB10 sulfhydryl oxidation was shown to be CHCHD4-dependent, it was hypothesize to be imported to the IMS via the oxidative folding MIA mechanism [46]. In comparison with NDUFB10, the *Chlamydomonas* NUOB10 and other vascular plant orthologs contain only the C-$(X)_{11}$-C motif (Figure 12). If the cysteines are vital for mitochondrial import via the conserved oxidative folding mechanism, we should expect cysteine-to-serine substitutions of this motif to abolish NUOB10 import into the *Chlamydomonas* mitochondria and subsequent assembly into complex I. On the contrary, we observed that the single and double substitutions still allowed for incorporation of the NUOB10 variants into complex I, resulting in the accumulation of a mature holoenzyme (Figure 15). This finding is in

accordance with the observations in the C107S NDUFB10 patient tissues, where complex I deficiency was not consistent across all tissues. For example, the skin fibroblasts appeared to express more of the C107S NDUFB10 variant than other tissues, enabling normal range of complex I activity [46]. These results indicate that the C107S substitution did not abolish NDUFB10's import into human mitochondria. Instead, the degree of complex I deficiency in different tissues was due to tissue-specific differential expression of this C107S variant. Our results further emphasize that both cysteine residues of the C-$(X)_{11}$-C motif are not strictly essential for NUOB10 import into the *Chlamydomonas* mitochondria.

One alternative explanation for the above-mentioned observations is that NUOB10/NDUFB10 is imported into the mitochondria in a CHCHD4-dependent but cysteine-independent manner. Such a phenomenon has been described for the IMS-localized mitochondrial protease Atp23 [113], wherein the disulfide bonds are required for protein folding and stability, instead of mitochondrial import [113]. Atp23 is imported via hydrophobic interactions with CHCHD4, even in the absence of all cysteine residues. Interestingly, the yeast Atp23 contains ten cysteines, including one possible C-$(X)_{11}$-C motif similar to NDUFB10. All ten cysteines are involved in disulfide bond formation in the IMS, although the identity of the disulfide bond-forming cysteine residues remains unknown. A second possibility is that NUOB10 could be imported into the IMS via an alternative import mechanism independent of the MIA machinery, and is instead involved in post-import interaction with CHCHD4. For instance, it could be localized to the mitochondria by virtue of unknown internal targeting sequence(s) present in the protein, as is the case for BCS1, a factor required for complex III maturation [114]. On the other hand,

MICU1, another non-canonical substrate of the MIA machinery, contains a mitochondrial targeting sequence and is imported through a CHCHD4-independent transport. Post-import, CHCHD4 interacts with its substrate MICU1, catalyzing intermolecular disulfide bond formation, which enables the assembly of MICU1 into the mitochondrial calcium uniporter complex [105].

Because the cysteine variants of NUOB10 are still incorporated into complex I, the subunit does not require disulfide bond formation at the C-$(X)_{11}$-C motif for import (Figure 15). However, significant decrease in rotenone-sensitive NADH: duroquinone oxidoreductase activity and lower accumulation of the mature complex is observed (Figure 15). Therefore, we propose that the cysteines, while not strictly required for mitochondrial import, may have roles in protein folding and stability, assembly into the membrane arm, and assisting the ubiquinone reduction, and/or proton-pumping capacity of complex I. The exact contribution of the C-$(X)_{11}$-C motif to complex I activity and assembly still remains to be determined.

In summary, we have successfully used *Chlamydomonas* for testing the impact of human pathogenic nuclear mutations on complex I assembly/activity, revealing the utility of a unicellular plant model as an experimental system of study for unraveling the molecular basis of complex I deficiencies. Even accounting for variations in the assembly process between *Chlamydomonas* and human complex I [28], it is reasonable to expect that analyzing nuclear pathogenic mutations in *Chlamydomonas* will provide insight into their consequence on complex I function. Our work opens up new avenues of exploration through a systematic approach, where substitutions of all conserved residues in complex I

subunits, individually and in concert, could be methodically employed to document the functional importance of each residue in complex I assembly and activity.

Chapter 3. AMC1, a large, low complexity protein is required for expression of mitochondrial-encoded *nd4*

This chapter has been previously published:

Subrahmanian, N, **Castonguay, AD**, Remacle, C, Hamel, PP. **(2020)** Assembly of Mitochondrial Complex I Requires the Low-Complexity Protein AMC1 in *Chlamydomonas reinhardtii*. *Genetics.* **214**: 895-911; *https://doi.org/10.1534/genetics.120.303029*

Contributions:

Castonguay, AD: Figure 21A and C, 24E, 25A-B, and editing/revision process.

3.1 Abstract

Complex I is the first enzyme involved in the mitochondrial electron transport chain. With > 40 subunits of dual genetic origin, the biogenesis of complex I is highly intricate and poorly understood. We used *Chlamydomonas reinhardtii* as a model system to reveal factors involved in complex I biogenesis. Two insertional mutants, displaying a complex I assembly defect characterized by the accumulation of a 700 kDa subcomplex, were analyzed. Genetic analyses showed these mutations were allelic and mapped to the gene *AMC1* (*Cre16.g688900*) encoding a low-complexity protein of unknown function. The complex I assembly and activity in the mutant was restored by complementation with the wild-type gene, confirming AMC1 is required for complex I biogenesis. The N-terminus of AMC1 targets a reporter protein to yeast mitochondria, implying that AMC1 resides and functions in the *Chlamydomonas* mitochondria. Accordingly, in both mutants, loss of AMC1 function results in decreased abundance of the mitochondrial *nd4* transcript, which encodes the ND4 membrane subunit of complex I. Loss of ND4 in a mitochondrial *nd4* mutant is characterized by a membrane arm assembly defect, similar to that exhibited by loss of AMC1. These results suggest AMC1 is required for the production of mitochondrially-encoded complex I subunits, specifically ND4. We discuss the possible modes of action of AMC1 in mitochondrial gene expression and complex I biogenesis.

3.2 Introduction

The mitochondria are important organelles for cellular energy production. Mitochondrial ATP synthesis is chemiosmotically coupled to the electron transfer chain (ETC), which is located in the mitochondrial inner membrane [31]. The ETC involves four major membrane bound complexes that mediate electron transfer from the substrates, NADH, or succinate, to the terminal electron acceptor O_2 [30]. Mitochondrial complex I (EC 1.6.5.3) is the first and largest enzyme of the ETC. This multimeric enzyme consists of one flavin mononucleotide molecule and eight iron-sulfur (Fe-S) clusters that catalyze electron transfer from NADH to ubiquinone [33, 115]. Conserved from bacteria to humans, complex I is a major entry point for electrons into the respiratory chain. For each molecule of NADH oxidized by complex I, four protons are pumped from the matrix to the intermembrane space, leading to the production of 2.7 ATP molecules per oxygen atom reduced [116, 117].

Electron micrography and X-ray crystallography of complex I from multiple organisms have revealed a L-shaped structure, with a hydrophobic arm embedded in the inner mitochondrial membrane and a hydrophilic arm protruding into the matrix (Figure 21b and Figure 24b) [55, 115, 118-121]. The hydrophilic arm contains domains for NADH and ubiquinone binding, whereas the hydrophobic arm is required for proton pumping. Eukaryotic complex I consists of > 40 nuclear- and mitochondria encoded subunits [28, 38]. Among these, 14 are considered “core” subunits as they are conserved in bacterial complex I and define the minimal requirement for catalytic activity. In addition, 30 or more “accessory” subunits are present, contributing to a total size of 1 MDa [28, 38]. Because of the large number of subunits of dual genetic origin, complex I assembly requires precise

coordination of various cellular processes: the transcription of genes encoding subunits, translation of the transcripts in the cytosol and the mitochondria, mitochondrial import of cytosol synthesized subunits, synthesis and insertion of the cofactors, and the correct folding of each subunit. Subsequently, the hydrophilic and hydrophobic subunits need to be assembled in a stepwise manner into individual modules, which are then combined to form the holoenzyme. This basic model of complex I biogenesis is common for all eukaryotes although the sequence of subunit addition in individual modules may vary according to the organism [104]. This elaborate process requires the assistance of nonstructural proteins that facilitate each step of biogenesis. Such proteins, referred to as biogenesis/assembly factors, are involved in the various stages of complex I biogenesis, but do not form part of the final holoenzyme [28, 41].

To date, 20 proteins have been associated with mammalian complex I biogenesis, of which 12 have been linked to human diseases [122-124]. Orthologs for most of the mammalian assembly factors are present in plants, although their function in plant complex I assembly is yet to be experimentally confirmed [28]. In *Arabidopsis*, a role in complex I assembly has been validated for only three factors: INDH, GLDH (reviewed in [28]), and CIAF1 [125]. INDH, an ortholog of the mammalian factor Ind1, was originally postulated as involved in Fe-S delivery to complex I subunits. However, in *Arabidopsis*, an alternate role in the translation of mitochondrially encoded complex I subunits was demonstrated for INDH [126, 127]. GLDH is a plant-specific assembly factor that localizes in the inner mitochondrial membrane and faces the intermembrane space. *Arabidopsis* GLDH is associated to three sequential, membrane-bound assembly intermediates [128, 129].

Finally, CIAF1 is a LYR domain containing mitochondrial matrix protein that interacts with the matrix arm 23 kDa TYKY subunit of complex I [125]. Loss of CIAF1 in *Arabidopsis* resulted in the accumulation of the 650 and 800 kDa assembly intermediates. A possible role in Fe-S cluster assembly was postulated based on the ability of CIAF1 to rescue a yeast mutant deficient for Fe-S cluster biogenesis.

Considering the high degree of intricacy in complex I biogenesis, it is reasonable to expect numerous factors remain to be discovered. Very little is known about factors specifically dedicated to the expression of mitochondrially-encoded complex I subunits, except in vascular plants [130]. To identify additional complex I biogenesis factors, we selected the unicellular photosynthetic alga, *Chlamydomonas reinhardtii*, as an experimental system [38]. The photosynthetic capacity of *Chlamydomonas* allows respiratory mutants to be viable [64]. This is in contrast to mammalian organisms which cannot survive when respiration is abolished. Additionally, the *Chlamydomonas* ETC contains type-II NADH dehydrogenases, absent in mammalian systems, which can partially compensate for loss of complex I activity [65] and allow the survival of complex I-null mutants under respiratory conditions. However, the alternate enzymes cannot contribute to the proton gradient, purportedly decreasing the efficiency of ATP synthesis. Hence, complex I-null mutants exhibit a respiratory growth defect that manifests as a Slow growth In the Dark (SID) phenotype [8, 22, 24, 27, 66]. Utilizing this SID phenotype, a number of mutants for mitochondria- and nuclear-encoded complex I subunits have been isolated in *Chlamydomonas* (reviewed by Salinas *et al* [29]), providing valuable insights into the sequence of complex I assembly in this organism. So far, only one biogenesis

factor, the NDUFAF3 ortholog, has been experimentally confirmed as required for *Chlamydomonas* complex I assembly [99].

In this study, we report a previously unidentified biogenesis factor, AMC1, required for the assembly of the distal membrane arm of complex I. We show that the N- terminus of AMC1 is capable of targeting a reporter protein to the mitochondria. Loss of AMC1 specifically affects the abundance of mitochondrial *nd4* transcript, implicating a role for AMC1 in mitochondrial gene expression, a poorly understood process in *Chlamydomonas*.

3.3 Materials and methods

3.3.1 Strains and culture conditions

Chlamydomonas strains were grown in liquid or solid medium: Tris-acetate-phosphate (TAP), with 20 mM Tris-base and 17 mM acetic acid, or Tris-acetate-phosphate supplemented with arginine (1.9 mM) (TARG) [1]. For some strains, TARG supplemented with 25 mg/ml hygromycin B, or TARG supplemented with 25 mg/ml paromomycin, was used. Algal strains were grown at 25^{o}C, in continuous light at 50 mmol/m^{2}/sec or in the dark.

Wild-type strains 3A+ (*mt*$^{+}$*; arg7-8*) and 4C- (*mt*$^{-}$*; arg7-8*), derived from genetic crosses of the CC124 background strain, were used for generating the insertional mutants (provided by Dr. Rochaix, University of Geneva, Switzerland). The insertional mutagenesis was conducted as described in Barbieri et al. (2011) [8]. The original complex I mutants *amc1*(*4C10*) strain (*mt*$^{+}$*; amc1-1; APHVII; arg7-8*), *amc5*(*87D3*) strain (*mt*$^{+}$*;*

nuob10::APHVIII; arg7-8) [8], *amc11*(*10G11*) strain (*mt*$^{-}$*; amc1-2; APHVII; arg7-8*) [72], and their derivatives were used. In addition, the following haploid progeny obtained via genetic crosses were also used: *amc11*(65) (*mt*$^{-}$*; amc1-2; APHVII*) derived from *10G11*, *amc1(2)* (*mt*$^{+}$*; amc1-1*), and *amc1(27)* (*mt*$^{-}$*; amc1-1; arg7-8*) derived from *4C10*. The *amc5* strain complemented with the wild-type *NUOB10* gene is referred to as [*amc5; NUOB10*]. The Δ*nd4* strain was generated from the *dum11* strain [131]. Generation of strains from genetic crosses is detailed below. The key strains used in this study and their corresponding reference numbers in the *Chlamydomonas* Resource Center are listed in Table S1 of Subrahmanian, *et al*. 2020 [132].

Saccharomyces cerevisiae strain JM45 Δ*coq3* (*MATa, his4-580, leu2-3,112, trp1-289, ura3-52, coq3::LEU2*) [133] was grown at 28^{o}C in liquid or solid medium, containing glucose as the fermentable substrate or ethanol as the respiratory substrate [68]. Strains were transformed by the one-step transformation method [86, 87].

Chemo-competent *Escherichia coli* DH5a strains were used for molecular cloning. *E. coli* was grown at 37^{o}C in Luria- Bertani broth and Luria-Bertani agar according to [69].

3.3.2 Genetic analysis

Genetic crosses were conducted according to [1]. Gametogenesis was induced by resuspending vegetative cells in TAP liquid medium lacking nitrogen (TAP-N), at 25°C in low light (0.5-1.0 μmol. $m^{-2}s^{-1}$), with shaking for 5 h. The gametes were mixed in equal proportions and incubated in light (50 μmol. m^{-2} s^{-1}) at 25°C overnight. To isolate meiotic

zygotes, the mixture was plated on TAP-N solid medium (containing 3% (w/v) select agar) and incubated in light (50 μmol $m^{-2}s^{-1}$) at 25°C for 5 days. The meiotic progeny was obtained through bulk germination on TARG solid medium, in continuous light at 25°C. To isolate vegetative diploids, the mating mixture was directly plated on selective medium mentioned below. Individual diploids and haploid progeny were sub-cloned to a single colony on solid medium and their mating type was determined by diagnostic PCR [74]. To determine the segregation of the SID phenotype in the *amc11* mutant, wild-type strain 1' (*mt*$^+$) [a 137C derivative, kindly provided by Dr. Claire Remacle, University of Liège, Belgium] was mated with *amc11(10G11)* (*mt*$^-$*; amc1-2; APHVII; arg7-8*). After bulk germination of the meiotic zygotes, haploid progeny were plated on TARG medium and the relevant phenotypes of the progeny were deduced by replica-plating. The haploid progeny derived from these crosses, *amc11*(65) (*mt*$^-$*; amc1-2; APHVII*), was retained for further analyses.

To test whether the *amc1* and *amc11* mutations are allelic, *amc11* x *amc1* diploids were generated by crossing *amc1(2)* (*mt*$^+$*; amc1-1*) with *amc11(10G11)* (*mt*$^-$*; amc1-2; APHVII; arg7-8*) with selection on TAP + HyB solid medium. In addition, haploid progeny from the *amc11* x *amc1* cross were obtained by crossing *amc1(2)* (*mt*$^+$*; amc1-1*) with *amc11(65)* (*mt*$^-$*; amc1-2; APHVII*). The meiotic zygotes were bulk-germinated and the haploid progeny (98 colonies) was tested for the relevant growth phenotypes by replica-plating.

3.3.3 Genomic DNA extraction and diagnostic PCR

Genomic DNA was extracted from *Chlamydomonas* by phenol-chloroform method [78]. The sequences of primers used in diagnostic PCRs and TAIL-PCRs are provided in Table S2 of Subrahmanian, *et al*. 2020 [132]. For diagnostic PCR analysis, GoTaq Polymerase (Promega, #M3008) was used according to manufacturer's protocol, with the addition of 2.5% (v/v) DMSO. Note that a denaturation temperature of 98ºC was used for *Chlamydomonas* genomic DNA / cDNA templates. The primers used for generating the amplicons in Figure 16D and the corresponding amplicon sizes are as follows: A [*Cre16.g688900* E2F33 / E2R9, 400 bp]; B [10G11 exon2F (1) / 10G11 AD1-F (2), 881 bp]; C [*au5.g6830* exon2F3 / 10G11 AD1-F (2), 469 bp]; D [10G11 exon2F (1) / amc1-insert-R2, 280 bp]; E [APH7F8 / 10G11 AD1-F (2), 592bp].

3.3.4 TAIL-PCR (Thermal Asymmetric Interlaced PCR)

TAIL-PCR was used to identify the sequence flanking the insertional cassette in the *amc11(10G11)* mutant as in [75]. The following partially degenerate primers were used for TAIL-PCR: AD1, AD2, and RMD228 [9, 75]. The iHyg3-specific primers, APH7F3, APH7F5, and APH7F8, were used to amplify the genomic DNA flanking the cassette at its 3′-end. The site of insertion was identified as exon 2 of gene *Cre16.g688900*, as provided by the *Chlamydomonas* genome database version 5.5 in Phytozome version 12 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii) [2]. The

sequence flanking the 5′-end of the cassette was obtained by conventional PCR, using the AMC1-specific primer au5.g6830 exon2F3 and the iHyg3-specific primer APH7R5 (Figure S1 of Subrahmanian, *et al*. 2020 [132]).

To identify the *amc1* mutation, diagnostic PCR was conducted and an insertional mutation was identified in exon 2. TAIL-PCR was used to identify the insertional sequence disrupting the *AMC1* gene in the *amc1-1* mutant by analyzing two independent *amc1-1* strains: *amc1(4C10)* and *amc1(27)*. Two sets of TAIL-PCR reactions were successful. In both, the partially degenerate primer RMD228 was used, in combination with the following *AMC1*-specific primers (for primary, secondary, and tertiary reactions, respectively): A) *Cre16.g688900* E2F32, *Cre16.g688900* E2F33, and *au5.g6830* exon2F2, and B) *Cre16.g688900* E2F33, *au5.g6830* exon2F2, and 10G11 exon2F (1). The complete sequence was retrieved by amplification with the primer pair amc1insert-F1/amc11del-R. The inserted DNA, disrupting the *AMC1* gene, was confirmed by sequencing two independent *amc1-1* strains: the original *amc1(4C10)* and its haploid progeny *amc1(27)* (Figure S2 of Subrahmanian, *et al*. 2020 [132]).

3.3.5 PCR-based screening of *Chlamydomonas* genomic library

An *ARG7*-based indexed cosmid library of *Chlamydomonas* genomic DNA, created by Dr. Jean David Rochaix's laboratory (University of Geneva, Switzerland) [76, 134], was screened for the presence of cosmids carrying the *AMC1* gene by diagnostic PCR. The *AMC1*-containing cosmid (referred to as 1H10) was identified using the primer pair

au5.g6830 exon2F3/10G11 AD1-F (2). In order to generate a construct containing only the full-length *AMC1* gene, the cosmid 1H10 was digested with *Bam*H1 restriction enzyme and re-circularized by T4 DNA ligase (Invitrogen, 15224041). The borders of *Chlamydomonas* genomic DNA insert, present in the original cosmid 1H10 (referred to as cosmid A) and the *Bam*HI digested and re-ligated cosmid (referred to as cosmid B) were sequenced to confirm the presence of the region carrying the gene of interest (Figure S3A of Subrahmanian, *et al*. 2020 [132]).

3.3.6 Biolistic transformation

Cosmids A and B were used for transformation by biolistics. The *ARG7* gene in the cosmids was used as a selection marker. The recipient arginine auxotrophic strain *amc11(10G11)* (*mt⁻; amc1-2; APHVII; arg7-8*), also referred to as the *amc1-2* mutant in this manuscript, was subjected to biolistic transformation using a homemade particle delivery device. The strains were grown in liquid TARG medium for 2-3 days until they reached exponential phase ($3 – 6 \times 10^6$ cells. mL^{-1}). The cells were plated on selective TAP medium at 10^8 cells/plate. For each bombardment, DNA was coated on sterile 0.6-0.9 μm tungsten particles (STREM Chemicals, #93-7437) by using 2 μg of cosmid, 16.7 mM Spermidine (Sigma, #S2626-5G), and 1 M $CaCl_2$. The bombardment was conducted at a helium pressure of 1.725 MPa and vacuum of -92 kPa. The plate was positioned 10.5 cm away from the nozzle containing the coated particles. The bombarded plates were first

incubated at low light overnight for recovery and then transferred to continuous light (50 μmol m^2s^{-1}).

Transformants from cosmid A were screened by replica-plating. Out of 311 transformants tested, 115 transformants displayed rescued growth in the dark. Transformants from cosmid B were screened by serial dilution. Out of 45 transformants, at least 40 displayed restoration of growth in the dark. Two transformants (two each for cosmids A and B), displaying rescued growth in the dark, were tested for biochemical activity and confirmed for complex I activity restoration. Out of the two, one transformant with cosmid A and one transformant with cosmid B were retained for further analysis.

3.3.7 Ten-fold dilution series

One loop of cells grown for 3–5 days on solid TARG plates were resuspended in 500 ml of liquid TARG medium. The cell density was measured spectrophotometrically at OD750 and diluted to an OD750 = 2.0. This normalized suspension was used as the starting material (1) for making five serial 10-fold dilutions (10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}). A volume of 8 μl for each dilution was plated on solid TARG plates. For scoring the SID phenotype, two plates were prepared simultaneously and incubated at 25^{o}C, one in continuous light and another in the dark, for at least 7 days. Ten-fold dilution series on yeast strains were conducted as described above, except that cells collected from solid medium were resuspended in sterile water. The cell density for yeast strains was measured spectrophotometrically at A600 and normalized to an OD600 = 1.0. The cells were plated

on two solid media and incubated at 28°C: single dropout synthetic medium lacking uracil (fermentable medium), and rich ethanol containing medium (respiratory medium).

3.3.8 Growth curves

Liquid cultures were inoculated with a starting cell density of 10^5 cells/ml in 50 ml TARG cultures. For each strain, three biological replicates were inoculated in continuous light at 25 mmol/m^2/sec and in the dark. Cell density was evaluated every 8 hr by measuring optical density at A750 [135], over a period of 10 days. Growth rate μwas calculated as 3.33 x ($\log_{10}N$ - $\log_{10}N_0$)/(t_N - t_0), where N is the final cell density at time t_N and N_0 is the initial cell count at time t_0. The generation time was calculated as 1/μ [1, 73]. The generation time for the strains in the light were determined between 36 and 108 hr, except for *amc1-2* (48–120 hr). The generation times for the strains in the dark were calculated between 36 and 156 hr, except for *amc1-2*, which was calculated between 72 and 240 hr. The wild-type and the *amc1-2* strain used were 4C- and *amc11*(*10G11*), respectively.

3.3.9 Enzymatic activity assays

Mitochondrial enzymatic activity measurements were conducted as described previously [24, 27, 81] with slight modifications. Cells grown for 2–3 days on solid medium were resuspended in MOPS-KOH extraction buffer (10mM MOPS-KOH, pH 7.4, 0.5 M mannitol, 100 mg/ml BSA, 0.5 mM PMSF). Cells were lysed by sonication using a

Branson Sonifier 150 (1/8 inch probe tip), at 12 W output for 2 x 30 sec. Following lysis, the extract was centrifuged at 480 3 x for 10 min, followed by 3000 x g for 5 min; the supernatant was centrifuged at 27,000 x g for 20 min, and the resulting pellet was the crude membrane extract. Complex I activity was determined as the rate of NADH oxidation, which was measured spectrophotometrically at A_{340}. The substrates used were 100 µM NADH (Amresco, #0384-1G) and 100 µM duroquinone (Sigma-Aldrich, #D22320-4). Specific activity was calculated using the molar extinction coefficient for NADH at ε_{340} nm = 6.22 mM^{-1} cm^{-1} in the absence and presence of 45 µM rotenone (MP Biomedicals, #150154), a complex-I-specific inhibitor. Complex II + III activity was determined as the rate of cytochrome *c* reduction and measured spectrophotometrically at A_{550}. It was conducted in the presence of 20.25 mM succinate (Acros Organics, #158751000), 1 mM KCN (Fisher Scientific, # P223I-100), and 56 µM equine heart cytochrome *c* (Sigma, #2506–500 mg). Complex II + III activity was calculated using molar extinction coefficient for cytochrome *c* at $\Delta\varepsilon_{550nm}$ = 19.6 mM^{-1} cm^{-1}, in the absence and presence of complex-III-specific inhibitor myxothiazol (3 µM) (Sigma, #T5580). Complex IV activity assay was conducted in the presence of 1%Triton X-100 and 56 µM reduced cytochrome *c*. Cytochrome *c* was reduced with two times the amount of sodium dithionite and purified with a PD10- desalting column with sephadex G-25 resin (GE Healthcare Life Sciences, #17085101) according to the manufacturer's protocol. Complex IV activity was calculated using molar extinction coefficient for cytochrome *c* at $\Delta\varepsilon_{550\ nm}$=19.6 mM^{-1} cm^{-1}, in the absence and presence of KCN (1 mM).

3.3.10 Blue-native polyacrylamide gel electrophoresis and in-gel activity

Partially purified membranes were extracted as described for enzymatic activity measurement. Complexes were separated by blue-native polyacrylamide gel electrophoresis (BN-PAGE) using 4–12% (w/v) acrylamide gradient gels as described [77]. Membranes were partially solubilized as described below. Membrane proteins (500 µg) were pelleted at 18,000 3 g for 20 min at 4^{o}C. The membranes were then resuspended in 180 ml of 2% (w/v) sodium n-dodecyl-b-D maltoside (Bioworld, #40430017-3) and incubated on wet ice for 1 hr. Unsolubilized membrane particles were removed by centrifugation at 18,000 3 g for 20 min at 4^{o}C. To the supernatant, 20 µl of 10% (w/v) sodium taurodeoxycholate hydrate (Sigma, #T-0875) was added [23]. Both n-dodecyl-b-D maltoside and taurodeoxycholate hydrate were dissolved in ACA buffer (750 mM aminocaproic acid, 0.5 mM EDTA, 50 mM Bis-Tris, pH 7.0). Partially solubilized membrane protein (200 µg) were loaded per lane. In-gel NADH dehydrogenase (complex I) activity was visualized as purple bands after incubating the gels in 100 mM MOPS buffer, pH 8.0, containing 1 mg/ml *p*-nitro blue tetrazolium chloride (Goldbio, #NBT2.5) and 0.2 mM NADH.

3.3.11 Immunoblotting analysis

For SDS-PAGE, 10 µg of proteins were separated by 12.5% acrylamide gel and immunoblotting was performed according to established protocols [78]. For BN-PAGE,

protein complexes were prepared and separated as described previously. Custom-made primary antibodies raised against *Chlamydomonas* complex I subunits 51 and 49 kDa, and TYKY subunits (referred to as NUO6, NUO7, and NUO8 in *Chlamydomonas*, respectively) were used [from Genescript, as described in Barbieri et al. (2011)]. Protein complexes separated via BN-PAGE were immunodetected with 1:3000 diluted a-51 kDa or 1: 3000 diluted a-49 kDa overnight at 4ºC. Protein extracts resolved via SDS-PAGE were immunodetected with 1:3000 diluted a-51 kDa, 1: 3000 diluted a-49 kDa, 1:2000 diluted a-TYKY, and 1:12,000 diluted a-cyt *f* for 3 hr at room temperature. The secondary antibody used was 1:10,000 diluted HRP-conjugated antirabbit goat antibody (Bio-Rad, #170-6515) and incubated at room temperature for 2 hr.

3.3.12 RNA extraction, RT-PCR, and qPCR

To determine the *AMC1* gene model (Figure 22), RNA was prepared using the Plant RNeasy Kit (Qiagen, #74904) with the following modifications. Cells were grown for 2-3 days on TARG solid medium. Cells (100 mg) were harvested and lysed by vortexing, with 9/10 volume of glass beads, for 5 min at 22ºC. RNA extraction was completed according to manufacturer's protocols. WT (4C-) RNA (2-5 μg) was treated with RQ1 RNase-free DNase I (Promega, #M6101). Reverse-transcription was achieved with 400 units of M-MLV reverse transcriptase (Life Technologies, #28025-013). Some regions of the cDNA could be amplified only if the first strand synthesis was conducted with Transcriptor High Fidelity cDNA synthesis kit (Roche #05081955001), using OligodT primers or random

hexamers as per manufacturer's protocol. Different regions of the *AMC1* cDNA were successfully amplified using either GoTaq Polymerase (Promega, #M3008) or Phusion High-Fidelity DNA Polymerase (NEB, #B0519S). The primer pairs used for successfully sequenced cDNA amplicons are provided in the legend of Figure 22. The amplicon generated using Cre16.g688900 5′UTR-F1/E2R10 was cloned into Promega pGEM-T easy Vector System. The resulting construct contains the sequence corresponding to the AMC1 N-terminal that was cloned in-frame to the *ubiG* reporter.

For real-time quantitative PCR (qPCR) of the mitochondrial transcripts, RNA was isolated by phenol-chloroform method modified from [82]. RNA was extracted from 2 x 10^8 cells grown in liquid culture. The cells were first resuspended in TEN buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 150 mM NaCl) and pelleted. The cells were resuspended in 150 µl of water and 300 µl of SDS-EB buffer (2% SDS, 400 mM NaCl, 40 mM EDTA, 100 mM Tris-HCl pH 8.0). Nucleic acids were extracted twice with equal volume of phenol-chloroform (pH 5.0) and RNA was precipitated with 1/3 volume of 8 M LiCl. RNA (8 µg) was treated with RQ1 RNase-free DNase I (Promega, #M6101). Reverse transcription was achieved with 800 units of M-MLV Reverse transcriptase (Life Technologies, #28025-013) using 1 µg of Random Hexamers (Promega, #C1181), according to the manufacturer's protocol. The absence of contaminating mtDNA was confirmed by diagnostic PCR of cDNA sample, across the two transcription start sites, with the primer pair ND5-EG-F2 and cox1-R (Table S2 of Subrahmanian, *et al*. 2020 [132]). The qPCR was conducted simultaneously on two dilutions of the cDNA and the results were compared for correlation. Amount of cDNA equivalent to 50 ng or 100 ng of total

input RNA was used as template for qPCR using SensiMix (Bioline, #QT-650-05) on a Mastercycler ep gradientS realplex thermocycler (Eppendorf). The qPCR reactions were denatured at 98ºC for 15 s, annealed at 60ºC for 20 s and extended at 72ºC for 20 s. Relative transcript levels were determined by normalizing the levels of target transcripts to the geometric mean of three reference transcripts, *TUA2*, *CBLP*, and *EIF1A*. For determining the relative mitochondrial DNA (mtDNA) content, 8 ng of total genomic DNA was used as template for qPCR. The mitochondrial *nd4* gene was used as the target gene and the nuclear *TUA2* gene was used as the reference. In all cases, normalization of target transcript or gene levels to the reference was performed by the Livak 2-ΔΔCt method [84]. Relative fold change for mtDNA content was calculated by normalizing to the average of the isogenic wild-type strain (set to 1.0). Relative transcript abundance is represented as percentage of wild-type strain. The amplification efficiency for each primer pair was determined by generating calibration curves for each pair of primers using two-fold dilutions of the template (wild-type cDNA). The efficiency was calculated from the slope of the calibration line as [10(-1/slope)-1] x 100. The efficiency and r^2 value from the calibration line are reported in Table S3 of Subrahmanian, *et al*. 2020 [132]. In Figure 24C, the *amc1-1* and *amc1-2* mtDNA content were normalized to their respective WT, 3A+ and 4C-, respectively. The strains tested in Figure 24D are WT (4C-), *amc1-2* [*amc11(10G11)*], and [*amc1-2; AMC1-B*]. The strains tested in Figure 24E are WT (3A+) and *amc1-1* [*amc1(4C10)*].

3.3.13 Plasmid construction to test for mitochondrial localization

The AMC1 N-terminal mitochondrial targeting sequence was heterologously expressed as a translational fusion with the bacterial UbiG protein in *S. cerevisiae*. For this purpose, plasmids pAHG (*CYC1* promoter + *ubiG*), pQMG (*CYC1* promoter + *ubiG* with 5'-end sequence encoding the COQ3 N-terminal mitochondrial targeting signal), and pAH3 (*CYC1* promoter + *COQ3*), provided by Dr. Catherine Clarke (University of California, Los Angeles) [136], were used. The sequence encoding the AMC1 N-terminus (1-59 amino acids) was amplified, using the primers PHA-F and PHA-R, from a plasmid containing a cloned fragment of the *AMC1* cDNA. The PCR product was cloned into the *Hind*III site of the plasmid pAHG [136], carrying the *ubiG* gene under the control of the constitutive *CYC1* promoter, by In-Fusion Cloning (In-Fusion HD cloning kit, Clontech, #639648). The resulting construct, containing the sequence corresponding to 1-59 amino acids of AMC1 at the 5'-end of the *ubiG* gene, is referred to as pAHG59. The pAHG, pQMG, pAH3, and pAHG59 plasmids, each containing the *URA3* selectable marker, were introduced into the Δ*coq3* strain by one-step transformation [86].

3.3.14 Data availability

Strains and plasmids are available upon request. The authors affirm that all data necessary for confirming the conclusions of the article are present within the article,

figures, tables, and supplemental information. Supplemental material available at figshare: https://doi.org/10.25386/genetics.11872362.

3.4 Results

3.4.1 Two complex I mutants harbor mutations in the same gene *AMC1*

A forward genetic screen was executed to isolate complex I mutants in *Chlamydomonas* [8, 72]. Insertional mutagenesis was conducted using the wild-type strain as a recipient with the iHyg3 cassette which encodes the *APHVII* transgene and confers hygromycin B resistance. In this study, we investigated two insertional *amc* (assembly of mitochondrial complex I) mutants, *amc1* [8] and *amc11* [72], that display a characteristic SID phenotype with severe complex I deficiency (Figure 16, A and B).

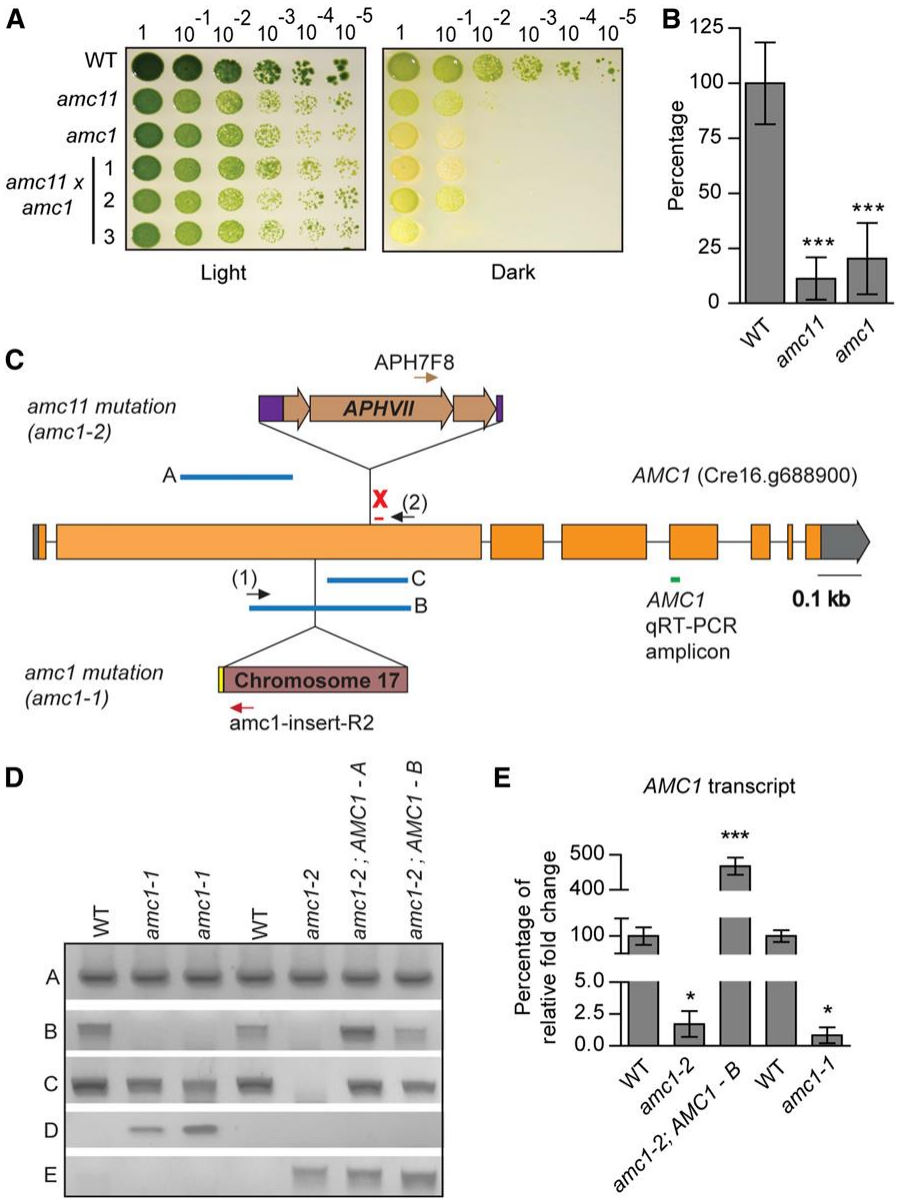


Figure 16 The insertional mutations in the *amc1* and *amc11* are allelic.

(A) The SID phenotype of the *amc1* and *amc11* mutants with three independent *amc1* x *amc11* diploids was tested by 10-fold dilution series. The dilutions were plated on acetate containing medium and incubated, in the light and in the dark, for 16 days. (B) Complex I (rotenone sensitive NADH: duroquinone oxidoreductase) activity was determined for the haploid strains shown in A from six biological replicates. The average complex I activity of WT (4C2) is 67.7 6 12.61 nmol NADH oxidized/min/mg protein. The activities are represented as percentage of WT. The significance of complex I deficiency with respect to WT was determined by two-tailed unequal variances t-test and the symbol *** $P < 0.001$. (C) The positions of the insertional mutations in the *Cre16.g688900* (*AMC1*) gene of the

amc1 and the *amc11* mutants are represented here. The gray rectangles indicate UTRs, the orange rectangles indicate the regions of the exons corresponding to the coding sequence, and the black lines correspond to the introns of the *AMC1* gene, as predicted by the *Chlamydomonas* genome database v5.5. In the *amc11* mutant, the insertional mutation is caused by the iHyg3 cassette. The red symbol "X" marks the location of the 42 bp deletion in *AMC1* of the *amc11* mutant. The purple rectangles in the inserted sequence represent the 568 and 42 bp sequences that cointegrated at the 59- and 39-ends of the cassette, respectively. In the *amc1* mutant, the insertional mutation is a 1.535 kb sequence of an intergenic region from chromosome 17. The yellow rectangle represents 28 bp of random DNA cointegrated at the 59-end of the insertion site. The blue lines indicate the genomic DNA regions (A–C) amplified by PCR (Figure 1D). The brown, red, and black arrows represent the primers specific for the iHyg3 cassette, chromosome 17 intergenic DNA, and *AMC1*, respectively, used to amplify amplicons D and E in panel D. The green line indicates the region of transcript amplified for quantitative RT-PCR in panel E. (D) The molecular lesions in *AMC1* were confirmed by diagnostic PCR using *AMC1*-specific primers. Two *amc1-1* strains, the original *amc1(4C10)* and a haploid progeny *amc1(27)*, and the *amc1-2* strain *amc11(10G11)* were used to confirm the mutation in *AMC1*. Amplicons D and E were amplified with primers indicated in panel C: D [10G11 exon2F (1)/amc1-insert-R2]; E [APH7F8/10G11 AD1-F (2)]. (E) Real-time quantitative PCR was used to assess the quantity of *AMC1* messenger RNA relative to three reference transcripts *TUA2*, *EIFA*, and *CBLP*. The average was obtained from three biological replicates, each including two technical replicates. The error bars represent SD of the mean. The results are represented as percentage of fold change relative to respective WT (3A+ for *amc1-1* and *4C2* for *amc1-2* and [*amc1-2; AMC1-B*]). The *amc1-1* and *amc1-2* strains tested here are *amc1(4C10)* and *amc11(10G11)*, respectively. Statistical significance was determined by two-tailed unequal variances t-test. * $P < 0.05$, *** $P < 0.001$. WT, wild type.

Previous analyses of *amc*1 had shown the SID phenotype did not co-segregate with the insertional cassette [8] and the *AMC1* locus had remained unidentified. Since the *amc11* mutant was also generated by insertional mutagenesis, we first tested if the complex I deficiency was linked to the insertional cassette via genetic analyses. Segregation of the SID phenotype and hygromycin B resistance (HyB^R) was monitored in the meiotic progeny obtained from the cross of *amc11* x wild type (Table 2). All the HyB^R meiotic progeny (total, 59) displayed a SID phenotype, while the hygromycin B sensitive (HyB^S) progeny (total, 96) displayed wild-type growth in the dark. Co-segregation of the SID phenotype with HyB^R indicates that the *amc11* mutation is tightly linked to the insertional cassette.

Table 1 The *amc11* mutation is linked to the insertional cassette

Parent strain mt^+		Parent strain mt^-		Meiotic progeny		
Strain	Phenotype	Strain	Phenotype	Phenotype	Total	SID
WT	arg^+, HyB^S	*amc11*	arg^-, HyB^R	arg^+, HyB^S	65	0
				arg^-, HyB^S	31	0
				arg^-, HyB^R	34	34
				arg^+, HyB^R	25	25
					155	59

Table 2 The *amc11* mutation is linked to the insertional cassette.

Meiotic progeny from the *amc11* x WT cross were plated on TARG medium and haploid progeny of all phenotypes were obtained by bulk germination. The relevant phenotypes were tested by replica-plating. WT and *amc11* strains used here are 19 strain and *10G11*, respectively. WT, wild type; arg+, arginine prototrophic strain; *arg2*, arginine auxotrophic strain; HyB^R, hygromycin B resistance; HyB^S, hygromycin B sensitivity.

To test if the *amc11* mutation is allelic to the previously isolated *amc* mutants [8], vegetative diploids were generated. Upon testing the respiratory growth phenotype, *amc11/amc1* diploids revealed a SID phenotype (Figure 16A). In addition, a total of 98 meiotic progeny obtained from an *amc1* x *amc11* cross displayed a SID phenotype. From these genetic analyses, we concluded that the mutations in *amc11* and *amc1* define the same genetic locus.

To map the mutations in the *amc11* and *amc1* strains, we conducted molecular analyses. The linkage between the insertional cassette and complex I deficiency in the *amc11* mutant can be attributed to the inactivation of a gene involved in complex I function by the insertional marker. Hence, we sought to retrieve the genomic sequence flanking the insertional cassette via TAIL-PCR. The mutation was mapped to exon 2 of the gene *Cre16.g688900* of unknown function (Figure 16C and [132]). A truncated form of the iHyg3 cassette, lacking the first 123 bp of the *TUB2* promoter, was inserted at this site. The insertional event was accompanied by a deletion of 42 bp of the genomic DNA at the site of insertion in the *Cre16.g688900* gene. Concomitant with the cassette insertion was the cointegration of foreign DNA (568 and 42 bp at the 59- and 39- ends of the cassette, respectively) corresponding to the herring sperm DNA used as a carrier during transformation [70]. Although herring sperm DNA sequence is currently unavailable, similarity searches using NCBI BLAST indicated the sequence integrated at the 59-end displays 94% identity to arctic cod fish DNA (*Boreogadus saida*).

In accordance with the TAIL-PCR analyses, DNA amplification of a region spanning the site of insertion in exon 2 (Figure 16D, amplicons B and C) failed to amplify in the *amc11* mutant, but was amplified in the wild type, confirming an insertion is present at this site. In addition, DNA was amplified, specifically in the *amc11* mutant and not in the wild type, at the 39-end junction between the cassette and the genomic region (Figure 16D, amplicon E). Therefore, the *amc11* mutation is caused by a 2.23 kb insertional sequence, consisting of iHyg3 cassette and herring sperm DNA, in exon 2 of *Cre16.g688900*.

Since the *amc1* mutation is allelic to the *amc11* mutation, we expected the *Cre16.g688900* gene to harbor a molecular lesion in the *amc1* strain. To determine if *amc1* is characterized by a mutation in this gene, we performed diagnostic PCRs on two *amc1* strains, across the full length of the *Cre16.g688900* gene. Similar to the *amc11* mutant, a region of exon 2 did not amplify in the *amc1* mutant strains (Figure 16D, amplicon B). The primer-binding sites for amplicon B are present in amplicons A and C, which are successfully amplified in the *amc1* mutant. Since we failed to detect a smaller amplicon B, we can rule out the occurrence of a major deletion. Therefore, the failure to detect amplicon B is suggestive of an insertion present in exon 2 of the *Cre16.g688900* gene in the *amc1* mutant. In addition, a region spanning the *amc11* insertion site amplified in the *amc1* mutant strains but failed to amplify in the *amc11* mutant, indicating that the insertion in *amc1* is upstream of the insertional site in *amc11* (Figure 16D, amplicon C). Since the insertional cassette is not linked with the *amc1* mutation, we hypothesized that the insertion in *amc1* is caused either by a truncated and nonfunctional cassette, the herring sperm DNA

used for mutagenesis, or fragments of *Chlamydomonas* genomic DNA. Insertional mutations caused by *Chlamydomonas* genomic DNA have been previously observed and are proposed to be caused by the integration of extracellular genomic DNA from lysed cells and uptaken during the transformation procedure [90]. To retrieve the insertional sequence in the *amc1* mutant, TAIL-PCR was performed with *Cre16.g688900*-specific primers. We discovered that the *amc1* mutation was caused by the insertion of a 1.535 kb section of *Chlamydomonas* genomic DNA, identical to an intergenic region from chromosome 17, in the reverse orientation (Figure 16C [132]). The insertion in *amc1* was accompanied by a 28 and a 10 bp insertion of random DNA at the 59- and 39-ends of the insertion site, respectively. The 28 bp sequence consists of 18 bp identical to an intron of an unrelated gene, *Cre16.g695800*. As predicted by diagnostic PCRs, the insertional site of the *amc1* mutation is 597 bp upstream of the *amc11* mutation. Further diagnostic PCRs conducted with *Cre16.g688900*-specific and inserted sequence-specific primers (Figure 16D, amplicon D) yielded amplicons of expected size only in the *amc1* mutants, confirming that the intergenic region from chromosome 17 is inserted in exon 2. Additional PCR analyses confirmed that there were no other major rearrangements present in the *Cre16.g688900* gene in the *amc1* or the *amc11* mutant strains. From these results, we conclude that the *amc1* and the *amc11* mutants carry insertional mutations mapping to the *Cre16.g688900* gene. Based on its role in complex I function, this gene shall henceforth be referred to as *AMC1*. In the rest of this manuscript, the allelic mutations in the *AMC1* gene present in the *amc1* and *amc11* strains shall be referred to as *amc1-1* and *amc1-2*, respectively.

To test if the insertional mutations in the *AMC1* coding sequence affected *AMC1* transcript levels, we conducted realtime quantitative RT-PCR analyses (Figure 16E). The *AMC1* transcript levels in the *amc1-1* and *amc1-2* mutants were significantly decreased (Figure 16E). The primers used for quantitative PCR analysis bind downstream of the insertion sites (green line in Figure 16C) and hence, the residual transcript levels detected could be caused by a chimeric transcript produced from the promoter elements present in the cassette or the intergenic region at the insertion site [83]. Since the *amc1-1* and *amc1-2* insertional mutations interrupt the coding sequence at ~3–4 kb from the start codon, we concluded that no full-length transcripts are produced in either mutant.

3.4.2 The *AMC1* gene restores complex I activity and assembly in the *amc1-2 mutant*

The *AMC1* gene does not encode for any known complex I subunits or assembly factors [28]. To test if the mutation in this gene is responsible for complex I deficiency, we complemented the *amc1-2* strain with the wild-type *AMC1* gene (Figure 17A). For this purpose, we transformed the *amc1-2* mutant with a cosmid (cosmid A) carrying the *AMC1* gene. The transformant, referred to as [*amc1-2; AMC1-A*], was restored for complex I activity (Figure 18A). In addition to the AMC1 gene, cosmid A contains two full-length genes (Figure 17A). To test that the *AMC1* gene is solely responsible for the restoration of complex I proficiency, a truncated form of the cosmid (referred to as cosmid B) containing only the full-length gene *AMC1* was generated (Figure 17A). The *amc1-2* strain transformed with cosmid B, referred to as [*amc1-2; AMC1-B*], also showed restoration of

complex I activity to wild-type levels, confirming that the disruption in the *AMC1* gene is responsible for the complex I defect (Figure 18A). As expected, complementation with the wild-type *AMC1* gene also resulted in restoration of the *AMC1* transcript levels in the [*amc1-2; AMC1-A*] and [*amc1-2; AMC1-B*] strains (Figure 16E and Figure 22). In fact, the *AMC1* transcript in [*amc1-2; AMC1-B*] showed a fivefold increase compared to the wild type (Figure 16E). Since the introduced *AMC1* gene is under the control of its native promoter, the increase in transcript levels is most likely due to the genomic location of the integrated wild-type gene, which we presume promotes increased *AMC1* expression in this transformant.

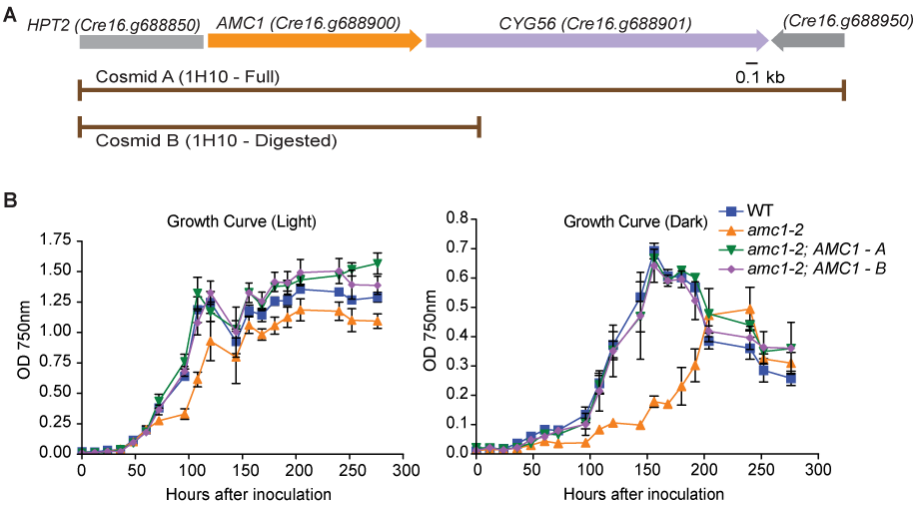


Figure 17 Rescue of the SID phenotype with cosmids containing the wild-type *AMC1* genomic DNA.

(A) Simplified map of the *Chlamydomonas* genomic DNA contained in the cosmids used for complementation studies. The 1H10 (Cosmid A) carries the truncated *HPT2* gene (*Cre16.g688850*), the *AMC1* gene (*Cre16.g688900*), the *CYG56* gene (*Cre16.g688901*), and *Cre16.g688950*. Cosmid 1H10 was digested by *Bam*HI restriction enzyme to obtain a truncated cosmid (Cosmid B) that only retains one full-length gene, namely *AMC1*. Both cosmids were used to transform the *amc1-2* mutant strain *amc11 (10G11)*. (B) The growth of WT, *amc1-2* [*amc11 (10G11)*], [*amc1-2; AMC1-A*], and [*amc1-2; AMC1-B*] was documented by measuring optical density at A750, in the light or in the dark, in the presence of acetate as a carbon source. The average of three biological replicates is reported here, with error bars indicating standard deviation of the mean. The generation time calculated from these growth curves is displayed in Figure 19A.

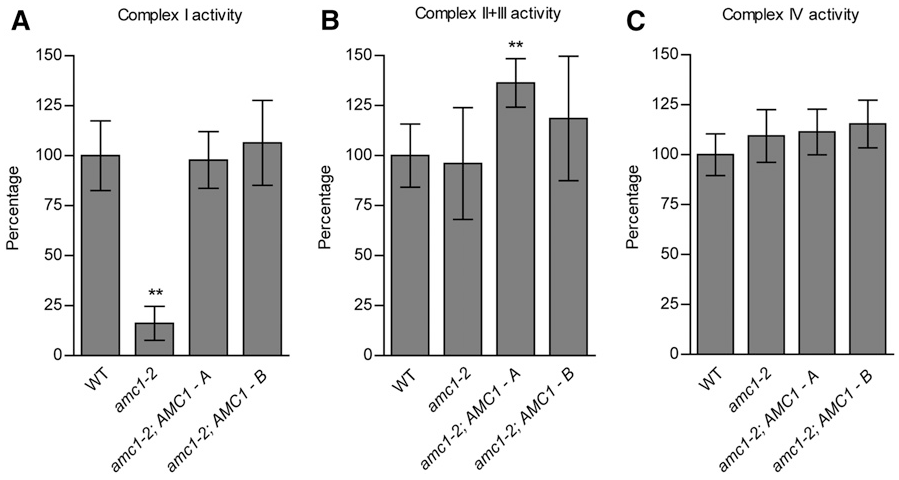


Figure 18 The *AMC1* gene restores complex I proficiency to the *amc1-2* mutant.

Enzymatic assays were conducted on partially purified membranes. The activities are displayed as percentage of the mean activity of wild type (WT), with the error bars indicating the percentage of SD relative to the mean. The WT and *amc1-2* strains used are 4C2 and *amc11(10G11)*, respectively. The *amc1-2* strain was transformed with cosmids containing the full length *AMC1* gene. The transformants containing the full cosmid (cosmid A) and the truncated cosmid (cosmid B) are referred to as [*amc1-2; AMC1 – A*] and [*amc1-2; AMC1 – B*], respectively. (A) Complex I (rotenone-sensitive NADH: duroquinone oxidoreductase) activity was determined from four biological replicates. The average complex I activity of WT is 69.28 6 12.08 nmol NADH oxidized/min/mg protein. Statistical significance was determined by two-tailed unequal variances t-test and the symbol ** ($P < 0.01$). The *amc1-2* mutant has significantly lower complex I activity compared to the WT ($P = 0.00122$). There is no significant difference between the activities measured for the *amc1-2; AMC1* and WT strains. However, there is a significant difference between (1) *amc1-2* and [*amc1-2; AMC1-A*] ($P = 0.00039$), and (2) *amc1-2* and [*amc1-2; AMC1-B*] ($P = 0.002545$). (B) Complex II + III (succinate: cytochrome *c* oxidoreductase) activity was determined from six independent biological replicates. The WT displayed an activity of 30.05 6 4.76 nmol cytochrome *c* reduced/min/mg protein. The *amc1-2* mutant displayed wild-type levels of complex II + III activity. The [*amc1-2; AMC1-A*] strain displays significantly higher activity ($P = 0.002$). (C) Complex IV (cytochrome *c* oxidase) activity was determined from four independent biological replicates. The WT displayed an activity of 248.52 +/- 25.8 nmol cytochrome *c* oxidized/min/mg protein.

To test if the mutation in the *AMC1* gene affects other respiratory complexes, complex II + III and complex IV activities were also determined. The *amc1-2* mutant and the corresponding complemented strains had wild-type levels of complex II + III and complex IV activities (Figure 18, B and C). We concluded that the *amc1-2* mutation, similar to the *amc1-1* mutation [8], only affects complex I.

Alongside the enzymatic analyses, we documented the effect of complementation on the growth phenotype of the *amc1-2* mutant by assessing growth in liquid and solid medium (Figure 19 and Figure 17). In accordance with the previously observed SID phenotype, the *amc1-2* mutant clearly displayed a slower growth rate in the dark, with a generation time of 45 hr, as opposed to the 28 hr for the wild-type strain. The [*amc1-2; AMC1-A*] and [*amc1-2; AMC1-B*] strains have a generation time of 24 hr in the dark, an indication that growth in respiratory conditions is fully restored (Figure 19A and Figure 17B). A rescue of the growth phenotype in the dark was also observed on solid medium (Figure 19B). Interestingly, the *amc1-2* mutant displayed a longer doubling time in the light (23 hr) compared to wild type (14 hr). This 1.6-fold slower growth rate in the light for the *amc1-2* mutant correlates with the smaller colonies observed during growth in the light on solid medium (Figure 19B). This phenotype has been previously observed for other complex I mutants as both photosynthesis and respiration contribute to the growth of cells in the light in acetate-containing medium [38]. The doubling time in the light was restored to wild-type levels upon complementation.

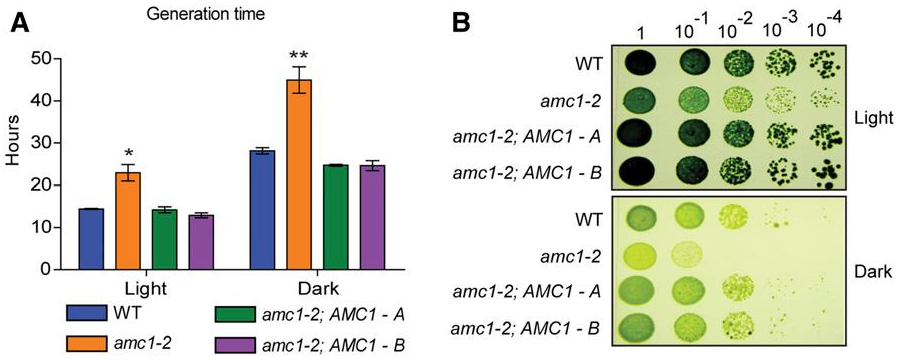


Figure 19 The SID phenotype of the *amc1-2* mutant is rescued by the WT-type *AMC1* gene.

The growth phenotypes of wild-type (WT) (4C2), *amc1-2* mutant, and the complemented strains [*amc1-2; AMC1-A*] and [*amc1-2; AMC1-B*], are compared here by studying growth on liquid and solid media. (A) The average generation time for each strain, determined from the growth curves in Figure 17B, is indicated here. The error bars represent SD of the mean. The two-tailed unequal variances t-test revealed a significant increase in the generation time of *amc1-2* with respect to WT, in the light and in the dark. In addition, the rescue of generation time in the dark was also significant for the complemented strains compared to the *amc1-2* mutant with a $P < 0.01$. * $P < 0.05$ and ** $P < 0.01$. (B) Restoration of the growth phenotype in the two [*amc1-2; AMC1*] complemented strains was confirmed by 10-fold dilution series, plated on acetate-containing medium and incubated in the light and in the dark for 12 days.

To determine the requirement of AMC1 for complex I assembly, mature complex I and partially assembled subcomplexes were separated by BN-PAGE. In-gel NBT staining assessing NADH dehydrogenase activity revealed mature complex I as a purple band at ~950 kDa in the wild type (Figure 21). Purple bands seen at higher molecular weights could be due to differential solubilization of membranes or association of complex I with other respiratory complexes such as complexes I + III_2 [22]. The accumulation of a 700 kDa subcomplex was observed in the *amc1-1* and the *amc1-2* mutants by in-gel staining and/or immunoblotting (Figure 21, A and C and Figure 20). However, the 700 kDa subcomplex is highly labile and its detection varies among independent membrane extractions [8]. The 700 kDa subcomplex also accumulates in the *nuob10-null* mutant (*amc5*), and the *nd4* and *nd5* mitochondrial mutants, because they lack membrane subunits required for the assembly of the distal membrane arm (Figure 21, A–C) [8, 22, 24, 63]. These results suggest that complex I membrane arm assembly was compromised in the *amc1-1* and *amc1-2* mutants. This assembly defect was corrected upon complementation with the wild-type *AMC1* gene (Figure 21). In addition to a subcomplex, lower levels of mature complex (at 950 kDa) can be detected in some extractions in the *amc1-1* mutant by in-gel staining and immunoblotting (Figure 21, A and C). Like the *amc1-1* mutant, traces of a 950 kDa complex was occasionally observed in the *amc1-2* mutant in certain membrane preparations (Figure 20). From these results, we conclude that *AMC1* plays a role in complex I membrane arm assembly.

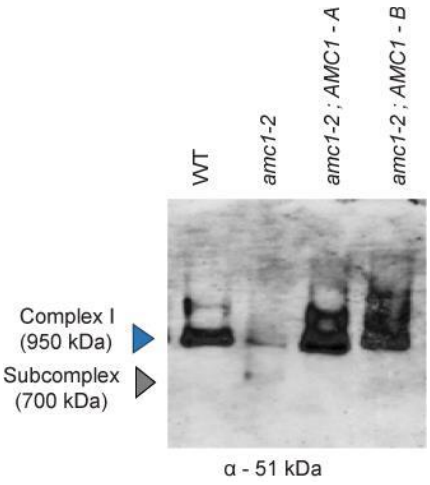


Figure 20 The *amc1-2* mutant display traces of a fully assembled complex.

Immunoblotting using anti-51 kDa antibody of complexes separated by BN-PAGE. Two hundred μg of protein were loaded per lane. While the labile 700 kDa subcomplex can always be detected in the *amc1-2* mutant, in occasional membrane extractions, low levels of a fully assembled complex I were also detected.

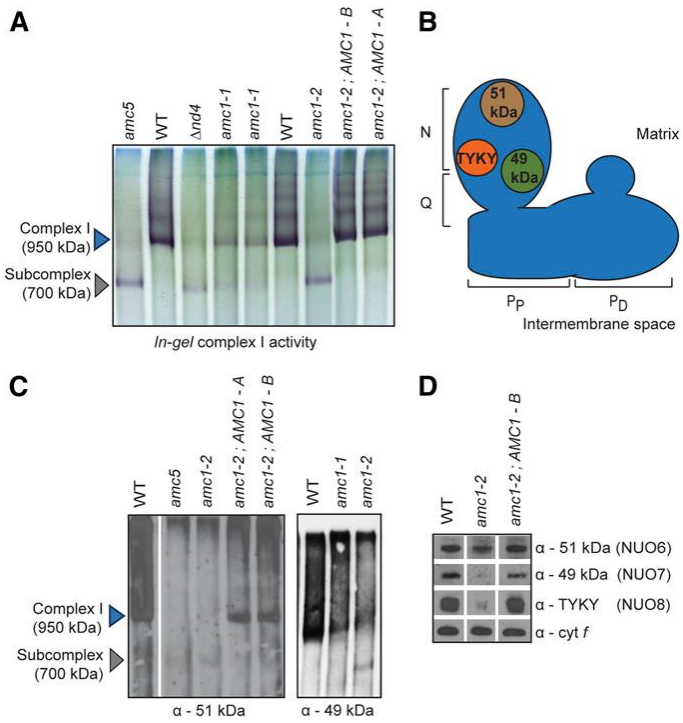


Figure 21 Complex I assembly is restored in the *amc1-2* mutant upon introduction of the *AMC1* gene.

(A) BN-PAGE was conducted on 200 mg of partially purified membrane fraction. In-gel complex I activity was detected by NBT staining. The *amc1-2* mutant accumulates a labile 700 kDa subcomplex, similar to the *amc5* mutant and Δ*nd4* mutant. The *amc1-1* mutant strains also accumulate trace levels of the 700 kDa subcomplex, in addition to a full complex. Two *amc1-1* strains, original *amc1(4C10)* (lane 4) and the haploid progeny *amc1(2)* (lane 5) were tested. (B) An illustration of complex I structure depicting the locations of the immunodetected subunits of the matrix arm, which consists of the NADH oxidizing N module and the quinone-binding Q module. The proton pumping membrane arm is divided into the proximal (P_P) and distal (P_D) modules. (C) Immunoblotting using α-51 kDa or α-49 kDa antibody, of complexes separated by BN-PAGE. Two hundred micrograms of protein were loaded per lane. The white vertical line depicts that the WT is from a different lane of the same immunoblot. The *amc1(2)* strain carrying *amc1-1* allele was tested in the right panel. (D) SDS-PAGE immunoblotting was conducted on 10 mg of partially purified membranes using polyclonal antibodies to detect soluble arm complex I subunits α-49 kDa, α-51 kDa, and α-TYKY. α -cyt*f* was used as a loading control. The white vertical lines indicate assembly of different lanes from the same immunoblot. In A, C, and D, one representative of three biological replicates is shown. The *amc1-2* strain *amc11(10G11)* and its corresponding complemented strains were tested here. WT, wild type.

Previously, it was observed that complex I deficiency results in decreased accumulation of complex I subunits [8, 99, 137]. Indeed, the *amc1-1* mutant displayed decreased steady-state levels of the 51 and 49 kDa, and TYKY subunits (referred to as NUO6, NUO7, and NUO8 in *Chlamydomonas*, respectively) [8]. Similar to the *amc1-1* mutant, the *amc1-2* mutant also accumulated lower levels of these three subunits [72], the levels of which were restored by complementation with the wild-type *AMC1* gene (Figure 21D). In summary, we have identified *AMC1* as a gene required for complex I assembly and activity.

3.4.3 The N terminus of AMC1 carries a mitochondrial targeting signal

According to the *Chlamydomonas* genome database JGI v5.5, the *AMC1* 8022 bp transcript specifies a large hypothetical protein of 2566 amino acids. No cloned complementary DNAs were available for *AMC1* to corroborate the predicted coding sequence [138]. To confirm the predicted coding sequence, RT-PCR was performed on total RNA isolated from the wild-type strain. Amplification of the *AMC1* complementary DNA was extremely challenging due to the low transcript abundance and high GC content. However, we were successful in sequencing 49.7% of the transcript (brown lines in Figure 22A). The retrieved sequences confirmed all the predicted exon-exon junctions, except for the exon 4/exon 5 junction that failed to amplify. The *Chlamydomonas* RNA-sequencing database (Oct. 2007 assembly) [2, 139, 140] corroborates the exon 4/exon 5 junction. From these results, we inferred that the coding sequence predicted by the JGI v5.5 is reliable.

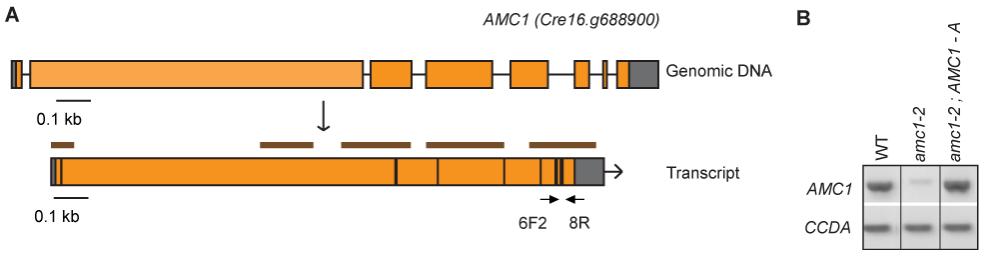


Figure 22 Transcript levels in the *amc1-2* mutant are restored by *AMC1* genomic DNA.

(A) The current *Chlamydomonas* genome database JGI v5.5 predicts the model for the *AMC1* gene as represented in the top panel. This gene model was experimentally corroborated by PCR amplification using *AMC1*-specific primers with total cellular cDNA prepared from the wild-type (4C-). The brown lines, in the bottom panel, identify the sections of the *AMC1* cDNA successfully amplified and confirmed by sequencing. The following primer pairs were successful in amplifying the sequence of the *AMC1* cDNA: Cre16.g688900 5′UTR-F1/E2R10; 10G11 exon2F (1) / amc11-del1R; Cre16.g688900 exon2F4/exon3R2; au5.g6830 exon3F / exon4R; Cre16.g688900 exon4F2 / exon4-3R; Cre16.g688900 exon5-F3 / exon-8R; Cre16.g688900 exon8-1F / Cre16.g688900 3′UTR-1R; Cre16.g688900 exon8-1F / 3′UTR-2R (Table S2 of Subrahmanian, *et al*. 2020 [132]). (B) Transcript abundance of *AMC1* was analyzed by RT-PCR in the wild-type (4C-), *amc1-2* [*amc11(10G11)*] strain, and the complemented strain [*amc1-2;AMC1-A*]. The *AMC1*-specific primers (au5.g6830 6F2 / Cre16.g688900 8R), used for amplification, are represented in (Figure 22A) with black arrows. *CCDA*, a gene involved in photosynthesis [141], was used as control for constitutive expression. Three independent biological replicates were performed, and one representative is shown in this figure. The vertical black lines indicate assembly of different lanes from the same gel.

The AMC1 protein sequence does not possess any recognizable domains indicating a possible function. If AMC1 plays a role in complex I biogenesis, it is highly plausible that mitochondrial localization is required for its function. Mitochondrial targeting is usually facilitated by a 10–80 amino-acid-long N-terminal targeting sequence which forms an amphiphilic a-helix [142, 143]. Algorithms have been designed to predict N-terminal mitochondrial targeting peptides from protein sequences. Mitochondrial targeting of AMC1 was predicted by the TargetP, MitoFates, and PredAlgo algorithms [144-147].

We experimentally tested whether the N-terminus of AMC1 can act as a mitochondrial targeting sequence by using a simple heterologous expression system with the UbiG reporter protein (Figure 23A). The *E. coli* UbiG and the *S. cerevisiae* Coq3 are proteins involved in ubiquinone biosynthesis. In *S. cerevisiae*, mitochondrial ubiquinone biosynthesis is essential for respiratory growth. It has been shown that *E. coli* UbiG can functionally complement the Δ*coq3* yeast mutant if UbiG is specifically targeted to the mitochondria with an appropriate targeting signal [136]. We utilized this heterologous system to test whether the AMC1 N-terminus can target *E. coli* UbiG to yeast mitochondria and restore respiration. For this purpose, a construct was created containing the sequence encoding the first 59 amino acids of AMC1, fused to the 59-end of the coding sequence of the ubiG gene (Figure 23B). The *ubiG* gene expressed without a mitochondrial targeting sequence failed to restore respiration [136].

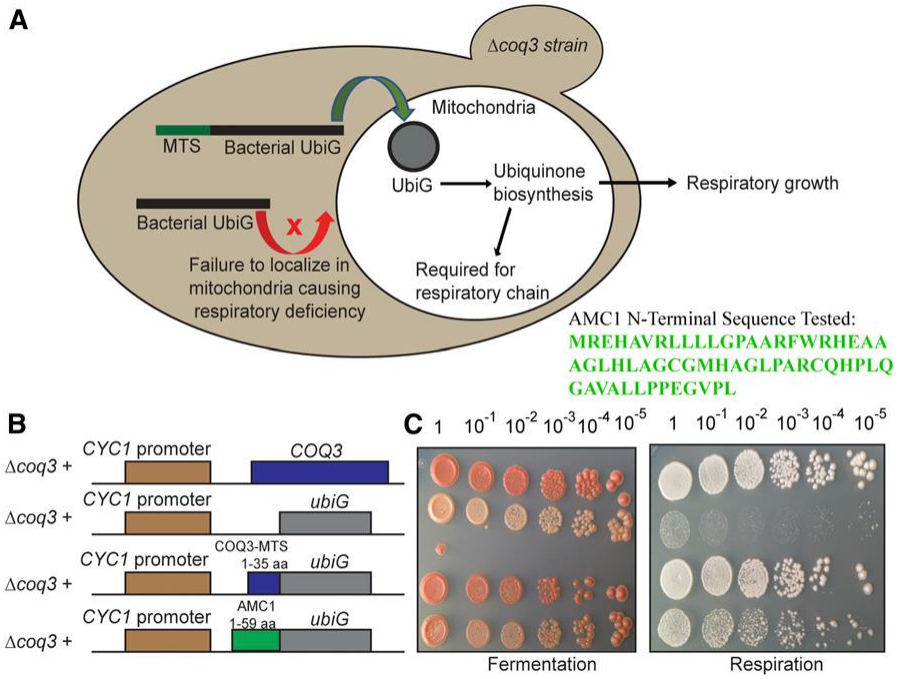


Figure 23 The N terminus of AMC1 targets the reporter protein UbiG to the yeast mitochondria.

(A) A simple model illustrating the heterologous system of expression for testing the mitochondrial targeting signal. In a *S. cerevisiae* Δ*coq3* mutant deficient for ubiquinone biosynthesis, production of the bacterial ortholog *UbiG* restores ubiquinone synthesis if it is specifically targeted to the mitochondria by virtue of a mitochondrial targeting signal (MTS). The N-terminal sequence of AMC1 that was tested is displayed in green font. (B) Simplified representation of the inserts carried by the different plasmids used. The construction of the plasmids is described in the methods. (C) The respiratory deficient Δ*coq3* mutant was transformed with constructs producing the UbiG reporter as a translational fusion with the first 59 aa in the N terminus of AMC1. The Δ*coq3* mutant transformed with the wild-type yeast *COQ3* gene, the *ubiG* gene lacking a MTS encoding sequence, and the *ubiG* gene with the sequence encoding the *COQ3* MTS were used as controls. Ten-fold serial dilutions were conducted on one representative transformant, plated on solid fermentable and respiratory media, and incubated at 28^{o}C for 7 days.

Accordingly, we observed significantly decreased respiratory growth of the Δ*coq3* strain expressing *ubiG*. On the contrary, *UbiG* fused with the *COQ3* mitochondrial targeting sequence or the first 59 amino acids of the AMC1 N-terminal sequence rescued respiratory growth of the Δ*coq3* mutant (Figure 23C). From these results, we conclude that the AMC1 N-terminal sequence is capable of targeting a reporter protein to the yeast mitochondria. Hence, AMC1 likely localizes in the *Chlamydomonas* mitochondria by virtue of its N-terminal sequence and functions in this compartment.

3.4.4 Loss of AMC1 results in decreased mitochondrial transcript abundance

Loss of AMC1 results in the accumulation of a 700 kDa membrane-bound subcomplex, a hallmark of mutants for the mitochondrial *nd4* and *nd5* genes whose products are part of the distal membrane arm [22, 63]. Hence, we hypothesized that AMC1 could control the mitochondrial synthesis of ND4 or ND5 subunits, and/or their assembly into the membrane arm. The mitochondrial genome of *Chlamydomonas* contains 13 intron-less genes, of which eight encode for proteins (Figure 24A): *cob* (apocytochrome *b*, subunit of complex III); *cox1* (subunit 1 of complex IV); *nd1*, *nd2*, *nd4*, *nd5*, *nd6* (subunits of complex I); and *rtl* (reverse-transcriptase like protein) [148]. The genome is transcribed in a bi-directional manner, generating the left and right transcriptional units (LTU and RTU, respectively) [149]. The transcription start sites are predicted to be located between the genes *nd5* and *cox1* [150]. The LTU is produced by the co-transcription of the *nd5*, *nd4*, and *cob* genes [149, 151], which are later processed to individual transcripts. The RTU is

generated by the co-transcription of the five protein-coding genes (*cox1*, *nd2*, *nd6*, *nd1*, and *rtl*), three transfer RNAs, and ribosomal RNAs (Figure 24, A and B), which are also processed into individual transcripts [152]. The mature transcripts begin at the start codon and do not have a 59-UTR [150]. This is similar to protein-encoding mitochondrial transcripts starting directly at the initiation codon in several organisms, such as humans [153] and fission yeast [154, 155], where the ribosomes display a clear preference for translation initiation on 59-UTR-less transcripts [156]. In contrast, the 59-UTR sequence plays an important role in the recognition of transcript by *S. cerevisiae* mitochondrial ribosomes [56]. A newly discovered feature of the *Chlamydomonas* mitochondrial transcripts is the presence of polycytidylation at the 39-end [150, 157].

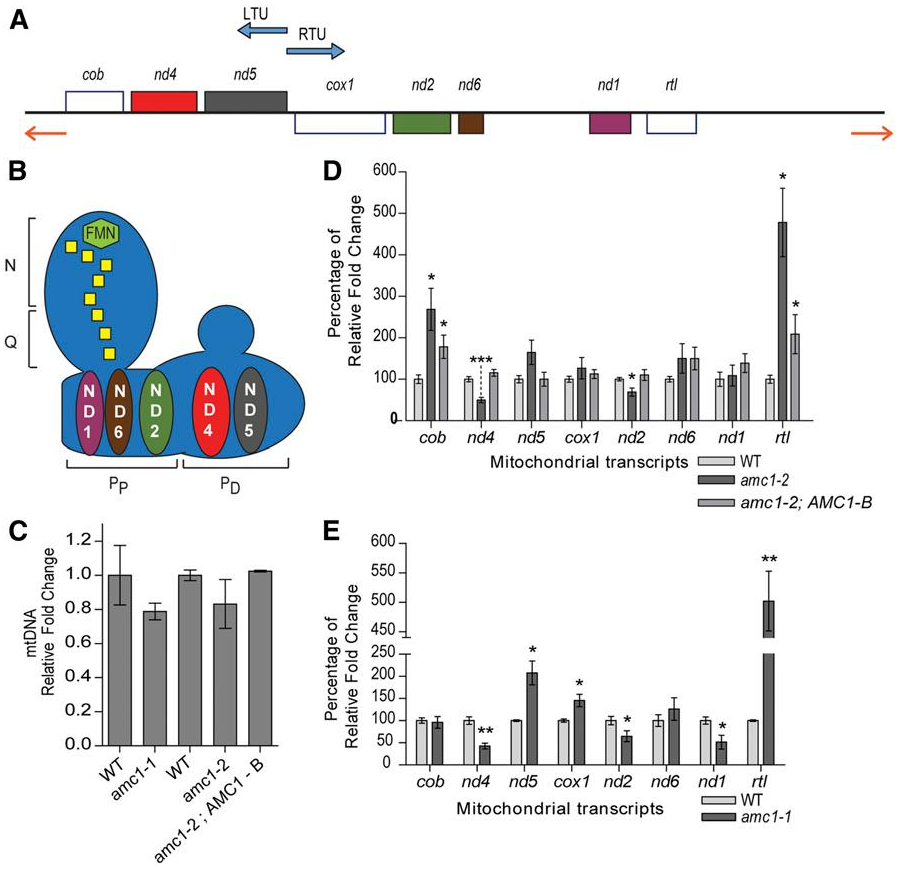


Figure 24 Loss of AMC1 results in decreased accumulation of the *nd4* transcript.

(A) Simplified map of the 15.8 kb mitochondrial genome of *C. reinhardtii*, adapted from [148]. The rectangles represent the eight protein-coding genes. The orange arrows represent the terminal inverted repeats. The blue arrows indicate the direction of transcription. The LTU contains *cob*, encoding the complex III subunit apocytochrome *b*; and *nd4* and *nd5*, encoding the complex I subunits ND4 and ND5, respectively. The RTU contains *cox1*, encoding subunit I of complex IV; *nd2*, *nd6*, and *nd1*, encoding the corresponding complex I subunits; and *rtl*, encoding a reverse-transcriptase like protein. (B) Simplified representation of L-shaped mitochondrial complex I embedded in the inner mitochondrial membrane. The flavin mononucleotide molecule is indicated in green and the Fe-S clusters are represented in yellow. The matrix arm consists of the NADH-oxidizing N module and the quinone-binding Q module. The proton-pumping membrane arm is divided into the proximal (P_P) and distal (P_D) modules. The putative locations of the mitochondrially encoded subunits of the membrane arm of complex I are represented here. (C) Real-time quantitative PCR was used to assess the relative quantity of mitochondrial DNA. The mitochondrial *nd4* gene was used as a target gene and the nuclear gene *TUA2*, encoding the alpha tubulin 2 protein, was used as the reference gene. The average was obtained from two biological replicates, each including three technical replicates, and the error bars represent SD of the mean. The results are represented as fold change relative to

WT (WT set to 1.0). The *amc1-1* strain tested in C and E was *amc1(4C10)* strain. The *amc1-2* strain tested in C and D was *amc11(10G11)* strain. (D and E) Real-time quantitative RT-PCR was used to assess the relative abundance of the mitochondrial protein-coding transcripts, which was determined with respect to the geometric mean of three reference messenger RNAs *TUA2*, *EIF1A*, and *CBLP*. The average was obtained from three biological replicates each analyzed in two technical replicates. The results are represented as percentage of fold change relative to WT (WT set to 100). Statistical significance was determined by two-tailed unequal variances t-test. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. WT, wild type.

First, we tested if the abundance of mitochondrial DNA was affected by the loss of AMC1. Real-time quantitative PCR analysis on purified genomic DNA demonstrated that there was no significant change in the mitochondrial DNA content in the *amc1-1* and *amc1-2* mutants, relative to wild type (Figure 24C).

To test if the complex I dysfunction by the loss of AMC1 is due to defects in the expression of the mitochondrial genes encoding complex I subunits, we assessed the abundance of the mitochondrial protein-coding transcripts by real-time quantitative RT-PCR analysis (Figure 24, D and E). The *amc1-1* and the *amc1-2* mutants displayed increased levels of the *nd5* transcript of the LTU compared to the wild-type (twofold and 1.6-fold for *amc1-1* and *amc1-2*, respectively). On the contrary, a twofold decrease in the *nd4* transcript was observed in both the mutants. Upon complementation with the wild-type *AMC1* gene, the *nd5* and *nd4* transcripts were restored to wild-type levels, indicating that the abundance of the LTU transcripts encoding ND5 and ND4 is dependent upon AMC1. Since the abundance of the *nd5* and *cob* transcripts of the LTU was not decreased, we can rule out an attenuation in the transcription of the LTU as the cause of the decrease in the *nd4* transcript levels.

In the RTU, while the *amc1-2* mutant displayed no significant difference in the *nd6* and *nd1* transcript levels, both the *amc1-1* and *amc1-2* mutants showed a 1.4- to 1.5-fold decrease in the *nd2* transcripts, respectively. In addition, the *amc1-1* mutant also displayed a twofold down-accumulation of the *nd1* transcript. Finally, the *rtl* transcript levels were increased by 4.8-fold and fivefold, in the *amc1-2* and *amc1-1* mutants, respectively. The *cox1* and *rtl* transcripts were not downaccumulated in the mutants, underscoring that the

transcription of the RTU is also not negatively affected. In summary, the *amc1-1* and *amc1-2* mutants displayed a decrease in the *nd4* and *nd2* transcript levels. The *amc1-1* mutant additionally displayed a decrease in the *nd1* transcript.

While our premise hypothesizes the decrease in the *nd* transcripts is due to loss of AMC1, it is possible that the mitochondrial transcript abundance is systematically affected because of complex I deficiency. To test this hypothesis, we assessed the mitochondrial transcript abundance in the *amc5* mutant (Figure 25). This mutant contains an insertional mutation in the nuclear *NUOB10* gene that encodes an accessory subunit located in the distal membrane arm of complex I [8, 72, 97]. The *amc5* mutant was chosen for comparison because it is another nuclear mutant that is characterized by a distal membrane arm assembly defect, similar to the *amc1-1* and *amc1-2* mutants (Figure 21A).

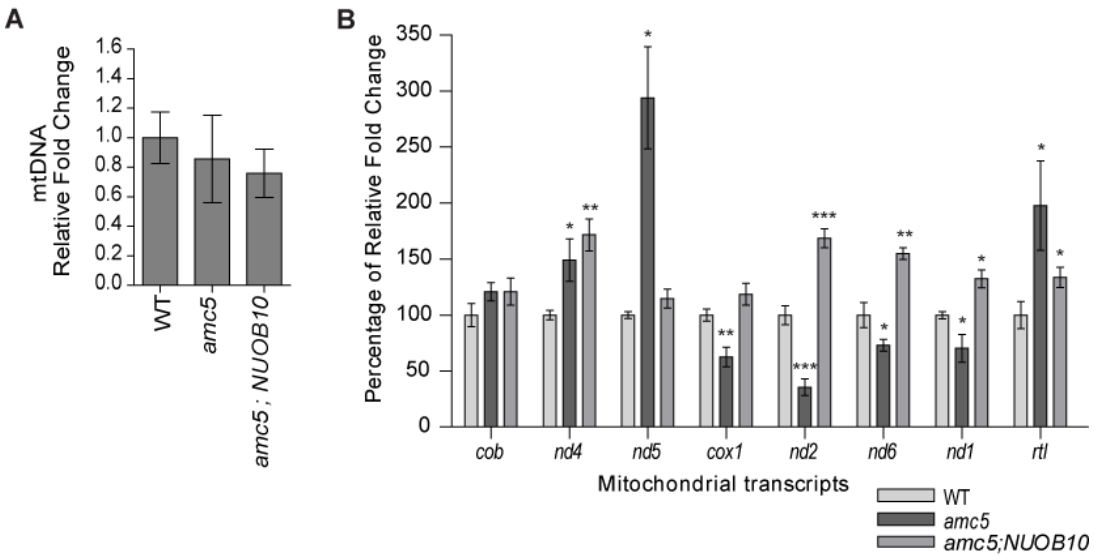


Figure 25 Mitochondrial transcript levels in a nuclear mutant characterized by the loss of the complex I subunit NUOB10.

(A) Real-time quantitative PCR (qPCR) was used to assess the relative quantity of mtDNA. The mitochondrial *nd4* gene was used as a target gene and the nuclear gene *TUA2,* encoding the alpha tubulin 2 protein, was used as the reference gene. The average was obtained from two biological replicates, each including three technical replicates, and the error bars represent standard deviation of the mean. The results are represented as fold change relative to WT (WT set to 1.0). (B) Real-time quantitative PCR (qPCR) was used to assess the relative abundance of the mitochondrial transcripts, *cob, nd4, nd5, cox1, nd2, nd6, nd1,* and *rtl.* The strains tested were WT (3A+), *amc5(87D3), and* [*amc5; NUOB10*]. The relative abundance of the mitochondrial transcripts was determined with respect to the geometric mean of the three reference transcripts *CBLP, TUA2,* and *EIFA.* The average is represented from three biological replicates, each analyzed in two technical replicates. The error bars represent standard deviation of the mean. The results are represented as percentage of fold change relative to WT (WT set to 100). The significance of difference in transcript abundance was determined with respect to the corresponding WT by two-tailed unequal variances *t*-test. The symbol * represents $p < 0.05$, ** represents $p < 0.01$, and *** represents $p < 0.001$.

In the *amc5* mutant, there was no decrease in the levels of the LTU transcripts. Interestingly, a 2.9-fold increase was observed for the *nd5* transcript, similar to the *amc1-1* and *amc1-2* mutants, which was restored to wild-type levels upon complementation with the wild-type *NUOB10* gene (Figure 25). Contrary to the *amc1-1* and *amc1-2* mutants, a 1.5-fold increase in the *nd4* transcript was observed in the *amc5* mutant.

In the RTU, the *rtl* transcript level was also increased in the *amc5* mutant similar to the *amc1-1* and *amc1-2* mutants. In addition, the three *nd* transcript levels (*nd2*, *nd6*, *nd1*) were decreased and were restored upon complementation. Th decrease in the *amc5* strain for the *nd2* and *nd1* transcripts is similar to the *amc1-1* and *amc1-2* mutants. However, based on their localization in the proximal membrane arm of complex I, complete loss of ND6, ND2, or ND1 is unlikely as this would result in a failure to assemble the 700 kDa subcomplex [24, 27].

Most importantly, the significant decrease in the *nd4* transcript, observed in both the *amc1-1* and *amc1-2* mutants, is not a feature of the *amc5* mutant. From these results, we can conclude that loss of AMC1 specifically affects the abundance of the *nd4* transcript. Previous studies on mitochondrial mutants of *Chlamydomonas* have revealed that loss of ND4 results in the accumulation of a 700 kDa subcomplex [27, 63], similar to the phenotype observed in the *amc1-1* and *amc1-2* mutants. Hence, we propose that the assembly defect resulting in the formation of the 700 kDa subcomplex in the *amc1-1* and the *amc1-2* mutants is due to a primary defect in the expression of the mitochondrial gene *nd4*.

3.5 Discussion

In this paper, we report the identification of a previously unrecognized player in *Chlamydomonas* complex I biogenesis. We have shown that (1) two allelic mutants exhibiting complex I deficiency, *amc1-1* and *amc1-2*, accumulate a 700 kDa subcomplex, diagnostic of an arrest in the assembly of the distal membrane arm of complex I; (2) both mutants carry insertional mutations in the unexplored *AMC1* gene; (3) complementation of the *amc1-2* mutant with the wildtype AMC1 gene restores assembly and activity, thereby confirming the role of AMC1 in complex I function; (4) the N terminus of AMC1 is capable of targeting a reporter protein to the mitochondria; and (5) loss of AMC1 results in decreased steady-state levels of the mitochondrial *nd4* transcript.

3.5.1 AMC1 is required for the assembly of the distal membrane arm of complex I

The multimeric complex I consists of three major functional modules: (1) the NADH-binding N module; (2) the quinone binding Q module, both part of the matrix-exposed peripheral arm; and (3) the proton-pumping P module of the membrane arm (Figure 21B and Figure 24B) [158]. The P module is further subdivided based on location with respect to the peripheral arm into the proximal (P_P) and the distal (P_D) ends of the membrane arm. Complex I biogenesis requires the synthesis of > 40 subunits followed by their precise step-wise integration into assembly intermediates to form the fully functional holoenzyme [104]. In instances where complex I biogenesis proceeds to form normal levels

of active holoenzyme, the assembly intermediates are normally so transient that they can hardly be detected by electrophoretic methods. The loss of a subunit or a biogenesis factor affects the assembly process, blocking the formation of the mature complex and resulting in the accumulation of assembly intermediates. The study of complex I mutants in various model systems has revealed the occurrence of specific soluble and membrane-bound intermediates stalled at particular stages of assembly, enabling to delineate the precise order of complex I assembly in each organism [28, 124, 158-160].

In this article we have reported two allelic insertional mutants, *amc1-1* and *amc1-2*, accumulating a ~700 kDa subcomplex (Figure 4). While multiple assembly intermediates have been identified in vascular plants and animals (reviewed in [28, 41, 124, 125]), only the ~700 kDa subcomplex has been well documented in *Chlamydomonas*. This assembly intermediate is a membrane bound subcomplex which has been previously observed in mutants lacking the subunits of the P_D module, such as ND4, ND5, and NUOB10 [8, 22, 24, 63, 99]. It can be detected by in-gel NADH/NBT staining, indicating that it consists of a functional peripheral arm capable of NADH dehydrogenase activity. Accordingly, proteomic analysis of the mitochondrial *nd5-null* mutant revealed the presence of at least 16 complex I subunits in the 700 kDa subcomplex [22]. This includes the seven core subunits from the N- and Q- modules required for NADH dehydrogenase activity, and the Pp module allowing the subcomplex to attach to the membrane. However, it lacks subunits of the P_D module required for a firm anchoring to the membrane, further highlighting this subcomplex does not contain the distal membrane arm of complex I. Since the 700 kDa subcomplex is loosely anchored to the mitochondrial inner membrane, it is

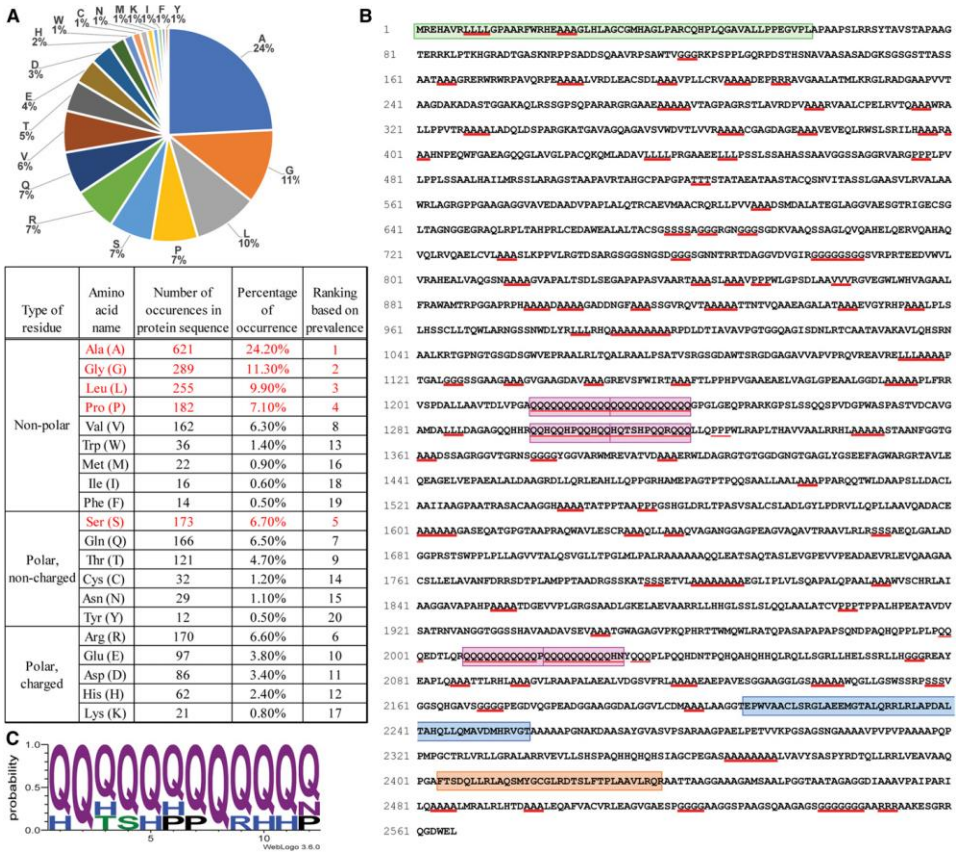
highly labile and its detection in membrane fractions varies with each extraction (Figure 21) [8]. In addition to the 700 kDa subcomplex, lower amounts of mature complex I are present in the *amc1-1* and *amc1-2* mutants indicating that limited membrane arm assembly can still occur in the absence of wild-type AMC1 (Figure 20 and Figure 21).

3.5.2 Structural features of the AMC1 protein

AMC1 is a large protein of 2566 residues with an expected molecular mass of 254 kDa. It is annotated in the *Chlamydomonas* genome database as a hypothetical protein (Cre16.g688900) and does not exhibit any significant sequence homology to proteins from bacteria, fungi, land plants, or animals. Heterologous expression in yeast has revealed that the first 59 amino acids of the AMC1 N-terminus are sufficient to target a functional reporter protein to yeast mitochondria (Figure 23). From these results, we infer that AMC1 localizes to the *Chlamydomonas* mitochondria where it functions in complex I biogenesis.

The AMC1 sequence is enriched for low-complexity regions, which are regions of biased amino acid composition. In AMC1, the predominant amino acids are Ala (24.2%), Gly (11.3%), Leu (9.9%), and Pro (7.1%) (Figure 26A). This feature is present in several nuclear-encoded Pentatrico Peptide Repeat (PPR) and Octotrico Peptide Repeat (OPR) proteins known to control the processing, splicing, editing, stabilization, and translation of transcripts in mitochondria and plastids [161-164]. PPRs and OPRs are defined by tandemly repeated motifs with degenerate amino acid sequences forming a-helical repeats that bind a target nucleotide sequence [165, 166]. However, PPR or OPR motifs are not

detected in AMC1. Yet, additional features similar to organellar RNA-interacting proteins are present in AMC1. First, like the chloroplast PPR and OPR proteins of *Chlamydomonas,* AMC1 contains stretches of amino acid repeats, predominantly alanine, glutamine, glycine, leucine, and proline (Figure 26B). Low-complexity proteins with polyalanine repeats have been well documented to have roles in transcriptional regulation, either through DNA- or RNA binding capacity [167]. In addition, six internally repeated polyQ (poly-glutamine) motifs with histidine and proline interruptions have been recognized by the MEME algorithm (Figure 26C) [168], and matches to multiple OPRs [164, 165, 169-171]. Interestingly, histidine and proline interruptions in the polyQ stretches were shown to decrease the tendency for polyQ motifs to aggregate by providing higher solubility [172]. PolyQ sequences are associated with various processes including transcriptional regulation and RNA binding [173].



Type of residue	Amino acid name	Number of occurences in protein sequence	Percentage of occurrence	Ranking based on prevalence
Non-polar	Ala (A)	621	24.20%	1
	Gly (G)	289	11.30%	2
	Leu (L)	255	9.90%	3
	Pro (P)	182	7.10%	4
	Val (V)	162	6.30%	8
	Trp (W)	36	1.40%	13
	Met (M)	22	0.90%	16
	Ile (I)	16	0.60%	18
	Phe (F)	14	0.50%	19
Polar, non-charged	Ser (S)	173	6.70%	5
	Gln (Q)	166	6.50%	7
	Thr (T)	121	4.70%	9
	Cys (C)	32	1.20%	14
	Asn (N)	29	1.10%	15
	Tyr (Y)	12	0.50%	20
Polar, charged	Arg (R)	170	6.60%	6
	Glu (E)	97	3.80%	10
	Asp (D)	86	3.40%	11
	His (H)	62	2.40%	12
	Lys (K)	21	0.80%	17

Figure 26 AMC1 is a low-complexity protein with polyQ regions.

(A) Amino acid composition of AMC1 is depicted as a percentage of total number of residues in the form of a pie chart for easy visual representation (top panel) and also as a table classifying the amino acids based on the properties of their side groups (bottom panel). The top five prevalent residues Ala, Gly, Leu, Pro, and Ser, are represented in red font in the table. (B) Amino acid sequence of AMC1. Any amino acid stretches that are three or more residues long, with polyA, polyQ, and polyG as the most prevalent, are indicated by a red underline. The MEME algorithm recognized a polyQ motif represented in pink boxes, the consensus sequence for which is shown in C. The green box indicates the first 59 amino acids fused to the *ubiG* reporter in Figure 23. The blue and orange boxes depict the sequences of the motifs 1 and 2, respectively, in AMC1 that were identified by MEME to have similarity within the C-termini of the *Volvox* and *Gonium* AMC1-like proteins. The consensus sequence for motifs 1 and 2 are provided in Figure 27. (C) WebLogo consensus sequence of the polyQ motif identified by the MEME program [168]. The letter height represents the probability of a particular amino acid at that position (y-axis) within the 12 amino acid repeats (x-axis).

Second, the large size of the protein (2566 residues) is a common feature of *Chlamydomonas* PPR and OPR proteins involved in chloroplast gene expression, the majority of which are > 1000 residues in length [161, 165]. Third, a-helical structures that are also a prominent feature in organellar RNA-interacting proteins [174], are predicted with high confidence and good correlation for AMC1 by Phyre2, Porter, and PSIPRED [175-177]. Finally, Phyre2, PSIPRED, and IUPRED also predicted multiple disordered regions, as expected from the presence of low-complexity regions in AMC1 [176-179]. Through their conformational variability and adaptability, disordered regions can support diverse functions such as scaffolding and recruitment of interacting partners [180]. From these similarities to proteins involved in the regulation of gene expression, we hypothesize that AMC1 may have a primary function in mitochondrial gene expression. We propose that AMC1 may directly bind RNA or may regulate gene expression by interacting with other RNA-binding proteins.

3.5.3 AMC1 is required for the expression of the mitochondrial *nd4* transcript

Analysis of mitochondrial transcripts has shown that loss of AMC1 results in a decreased accumulation of the *nd4* transcript, which was restored upon complementation (Figure 24). However, this defect was not observed in the *nuob10-null* mutant (Figure 25) even though it was also deficient in the assembly of the distal membrane arm (Figure 21A). Since decreased levels of *nd4* transcript were specifically reproduced in both the *amc1-1*

and *amc1-2* mutants, it is most likely that AMC1 functions in complex I biogenesis through its involvement in *nd4* gene expression.

The *nd4* transcript levels could be diminished due to (1) decreased transcription, and/or (2) increased degradation of the *nd4* transcript. The *nd4* transcript is synthesized as part of the polycistronic LTU and then processed to form the individual transcript for translation [149]. However, the levels of the *nd5* and *cob* transcripts, which are also part of the LTU, are not decreased in *amc1-1* and *amc1-2* (Figure 24C). Hence, we conclude that the problem does not lie with the transcription of the *nd4* gene.

A second possibility is that while the LTU is transcribed normally, it is not processed to release the individual transcripts for translation. However, we do not observe any associated decrease in *cob* and *nd5* transcript levels due to loss of processing or a resultant complex III deficiency in the *amc1-1* or *amc1-2* mutants (Figure 18B, [8, 72]). This indicates that *cob* is translated and thereby processed normally, and we can conclude that processing of the LTU is likely not affected.

A simple hypothesis to account for the assembly defect in the distal arm is a loss of ND4 and/or ND5. Since the abundance of the *nd4* transcript is reduced but not that of *nd5*, we hypothesize that the AMC1-dependent assembly defect is primarily due to loss of ND4. It is plausible that AMC1 functions in stabilizing the *nd4* transcript based on the fact that in the *amc1-1* and *amc1-2* mutants the *nd4* transcript accumulates to ~42 and ~50% of wild-type, respectively. However, organellar transcript levels may not be rate-limiting. In the chloroplast, transcripts can be reduced to < 10% of wild type by chemical treatment and the rate of translation still remains largely unaffected [181]. In the *Chlamydomonas*

mitochondria, a single deletion in the 39-UTR of the nd5 transcript still allows for the accumulation of the mature complex to 17% of wild-type levels, despite a drastic decrease in the mature *nd5* transcript levels [24]. Although the stability of the *nd5* transcript was severely affected by the 39-UTR mutation, a 700 kDa subcomplex was not detected in this mutant, indicating that the synthesis of ND5 and the assembly of the distal membrane arm were not significantly affected to result in the accumulation of the 700 kDa subcomplex [24]. In the *amc1-1* and *amc1-2* mutants, the *nd4* transcript abundance is reduced only by twofold but a 700 kDa subcomplex is detected, implying that the transcript levels may not be rate-limiting. We favor a model in which the primary defect due to loss of AMC1 function lies in the synthesis and assembly of ND4 into the distal arm. It must be noted that a complete lack of ND4 due to the loss of AMC1 is unlikely as traces of mature complex I are formed, in addition to the 700 kDa subcomplex (Figure 20 and Figure 21). Hence, we propose that loss of AMC1 function severely affects, but does not completely abolish, ND4 synthesis.

At the present time, the mode of action of AMC1 is speculative, but based on our results, we favor the hypothesis that AMC1 is primarily involved in translation of the *nd4* transcript. To account for the down-accumulation of the *nd4* transcript in the *amc1-1* and *amc1-2* mutants, we hypothesize that loss of ND4 synthesis may have a secondary effect on the *nd4* transcript stability, resulting in twofold decrease of the relative *nd4* transcript levels. However, it is also conceivable that AMC1 has dual function and controls both the stability and the translation of the *nd4* transcript. Based on the features of the protein sequence, the function(s) of AMC1 may be executed either through direct RNA binding or

through interactions with other RNA-binding proteins. Future investigations would involve documenting the effect of loss of AMC1 on mitochondrial translation. However, the tools currently available for investigating mitochondrial translation in *Chlamydomonas* are limited. Although a preliminary *in organello* translation study was reported, the pattern of migration of all the labeled polypeptides could not be definitively correlated with the expected gene products [182]. In addition, antibodies to detect the steady-state abundance of *Chlamydomonas* mitochondrial ND proteins are not available and our several attempts to detect the ND4 protein with antibodies against *Arabidopsis* Nad4 (PhytoAB Inc.) and *Polytomella* Nad4 (a generous gift from D. Gonzalez-Halphen) were unsuccessful (data not shown).

The ND4 subunit, consisting of 443 residues that form 14 conserved transmembrane helices, is involved in proton translocation [183, 184]. Mitochondrially-encoded proteins are known to be co-translationally inserted into the inner mitochondrial membrane [185]. Proteins assisting this process in *Chlamydomonas* remain unidentified. At this point, we cannot rule out that AMC1 may be involved in co-translational insertion of ND4 in the membrane or post-translational chaperoning of ND4 to facilitate its integration in complex I.

A defect in membrane arm assembly due to loss of mitochondrially-encoded ND4 has also been reported in plants and human cells [186, 187]. Although it is one of the core subunits, *Chlamydomonas* ND4 has only 27 and 39% sequence identity with mammalian and tobacco ND4, respectively, indicating a higher evolutionary distance between ND4 of various species [188, 189]. Such evolutionary changes in the protein sequence would

necessitate coevolution of the factors responsible for regulating the expression of the corresponding mitochondrial transcript [190]. Hence, it is conceivable that proteins involved in *Chlamydomonas* mitochondrial gene expression have also accrued modifications and features to accommodate the corresponding changes in their target transcripts. Consistent with this view is the fact that AMC1-like proteins can only be detected in other green algae such as *Volvox carteri* and *Gonium pectorale*, which are close relatives of *Chlamydomonas* (Figure 27). Interestingly, the AMC1-like proteins from *Volvox* and *Gonium* are also low-complexity proteins with polyG, polyS, and polyP stretches. In addition, the *Volvox* protein harbors two short polyQ motifs. Note that the similarity is limited to the C-terminus, an indication that this domain is probably important for the function of the AMC1-like proteins. A common feature of the three proteins is the presence of two partially degenerated motifs separated by 137–214 residues at the C-terminal end (Figure 26 and Figure 27). The two motifs are 48 and 34 residues in length. A common C-terminal motif in three algal counterparts might be indicative of a conserved algal domain with a common function. Further experiments are required to elaborate the mechanism of action of AMC1.

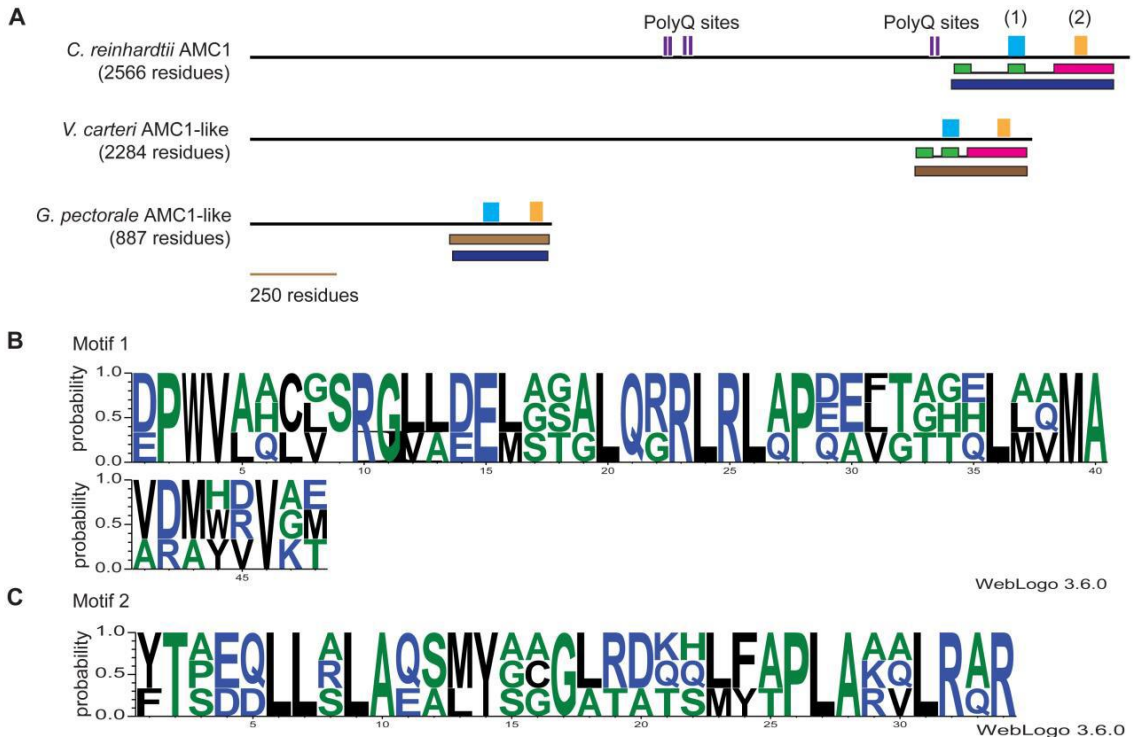


Figure 27 The C-terminus of AMC1 is similar to that of hypothetical algal proteins in *Volvox carteri* and *Gonium pectorale*.

Searches for AMC1-like proteins, via BLASTp (ALTSCHUL *et al.* 1990) and DELTA-BLAST (BORATYN *et al.* 2012), revealed that the AMC1 C-terminus (202 residues) has 37% identity with the C-terminus of another large hypothetical protein (2284 aa) from a closely related chlorophycean green alga *Volvox carteri* (Phytozome 12 Volvox database v2.1 ID: Vocar.0023s0109.1 / NCBI Accession: XP_002952084.1). AMC1 and the hypothetical *Volvox* protein show similarity at the C-terminus with another putative protein (887 aa in length) from the green alga *Gonium pectorale* (GPECTOR_23g121 / NCBI accession KXZ49034). While AMC1 shows only a weak similarity (29% identity amongst 432 residues), the *Volvox* hypothetical protein shows a significant similarity (36% identity amongst 339 residues) with the C-terminal region of the *Gonium* hypothetical protein. (A) Conserved C-terminal motifs and polyQ repeats in the AMC1 protein sequence are shown. The polyQ repeats shown in Figure 7 are represented here as purple bars. Similarity searches with AMC1 identified similarity in the C-terminal end within a *Volvox carteri* hypothetical protein. The pink box depicts 46% similarity in a 202 amino acid region, whereas the green boxes depict 64-68% in a 45 amino acid region. Subsequent similarity searches using the *Volvox* AMC1-like protein yielded a *Gonium pectorale* hypothetical protein. The brown box depicts a 296 amino acid region with 46% similarity to *Gonium* hypothetical protein. The dark blue box depicts the region of low similarity (36%) between the *Chlamydomonas* AMC1 protein and the *Gonium* hypothetical protein. Comparison of these three proteins using the MEME program [168, 191] identified the presence of two

different motifs (motif 1 and motif 2, represented by a light blue box and a yellow box, respectively) that is common to the C-terminus of all three proteins. The sequence of these motifs in AMC1 are represented in Figure 26B. (B) and (C) are the MEME-generated consensus sequences for motif 1 and motif 2, respectively [168]. Motifs 1 and 2 are predicted to be 48 and 34 residues in length, respectively. The height of each letter represents the frequency of occurrence of the corresponding amino acid (y-axis) at a given position in the motif (x-axis). The *E-value* for the occurrence of motif 1 and motif 2 are $1.3e^{-006}$ and $1.8e^{-004}$, respectively.

In conclusion, we have utilized a forward genetic screen in *C. reinhardtii* to reveal a novel nuclear-encoded factor involved in mitochondrial gene expression and required for complex I biogenesis. This study opens new avenues of exploration to delineate the processes of mitochondrial gene expression and their role in mitochondrial function.

Chapter 4. *LCLA1* encodes large protein participating in algal lipid metabolism

Excerpts of this chapter have been previously published:

Subrahmanian, N, **Castonguay, AD**, Fatnes, TA, Hamel, PP. **(2020)** *Chlamydomonas reinhardtii* as a plant model system to study mitochondrial complex I dysfunction. *Plant Direct.* **4**: 1– 16; *https://doi.org/10.1002/pld3.200*

Contributions:
Castonguay, AD: wrote this chapter and performed all experiments for figures except 44C (Ticoras, MT).

4.1 Abstract

The unicellular, photosynthetic, alga, *Chlamydomonas reinhardtii* has emerged as a popular model organism for the study of plant metabolism. *Chlamydomonas* can grow in the light fixing CO_2 via photosynthesis, mixotrophically provided acetate as an organic carbon source, or in the dark provided acetate via respiration. Multiple mutagenic screens for deficient growth in the dark have yielded mutants in mitochondrial respiration and acetate assimilation. One insertional mutant obtained from a previously conducted forward genetic screen, *lcla1,* also grew more slowly in photosynthetic conditions and was sensitive to high light, resembling the isocitrate lyase (*icl*) mutant which is unable to assimilate acetate. ICL catalyzes the first step of the glyoxylate bypass, a shunt of the tricarboxylic acid (TCA) cycle predominantly housed in the peroxisome. Introduction of a mutant allele of *ARG7* blocking the L-arginine biosynthetic pathway, another shunt of the TCA cycle, exacerbates the high light sensitivity of *lcla1* and *icl*. Since lipids are elevated in the *icl* mutant and high light is a stress condition known to induce plastid lipid synthesis, triacylglycerol (TAGs) content was measured in *lcla1*. As expected, TAGs containing lipid bodies were increased in the *lcla1* mutant due to a defect in lipid catabolism. Furthermore, *lcla1* and *icl* were resistant to an inhibitor of fatty acid synthesis in high light suggesting that the light sensitivity was due to lipotoxicity. Interestingly, mitochondrial and chloroplast morphology were abnormal in the *lcla1* mutant, possibly due to aberrant lipid homeostasis. The insertional cassette in *lcla1* was mapped to *Cre07.g329861*, a novel gene encoding a large protein with numerous low complexity regions but without any clearly identifiable features suggestive of a particular biological role and the gene renamed *LCLA1*

for Low Complexity protein in Lipid Accumulation. Further sequence analysis identified a putative peroxisome transmembrane domain at the N-terminus and a chimeric fusion protein containing this motif is sufficient to direct the reporter to the ER, the site of *de novo* peroxisome biogenesis in mixotrophic conditions. Shifting cultures to respiratory conditions, the reporter redistributes, matching that of peroxisome-targeted, fluorescent markers. While the precise function of LCLA1 remains to be elucidated, possible functions of LCLA1 are discussed.

4.2 Introduction

Chlamydomonas reinhardtii, a unicellular photosynthetic microalga (hereafter *Chlamydomonas*) is an established model organism for the study of photosynthesis, ciliopathies, mitochondrial respiration, and for production of economically important compounds such as biopharmaceuticals and biodiesel precursors [2, 3, 61, 192]. The popularity of *Chlamydomonas* stems from its ease of culture, genetic tractability, transgenic capability, rapid generation time (in comparison to land plants), *etc.* (for a more complete list see [1]). Indeed, *Chlamydomonas* only requires water, trace metals/minerals, air (O_2 and CO_2), light, and a nitrogen source such as ammonium to grow (Figure 28A) [1]. Differences in light condition and presence of supplemented carbon sources, mainly acetate, allow dissection of different biological pathways contributing to growth (Figure 28B) [1].

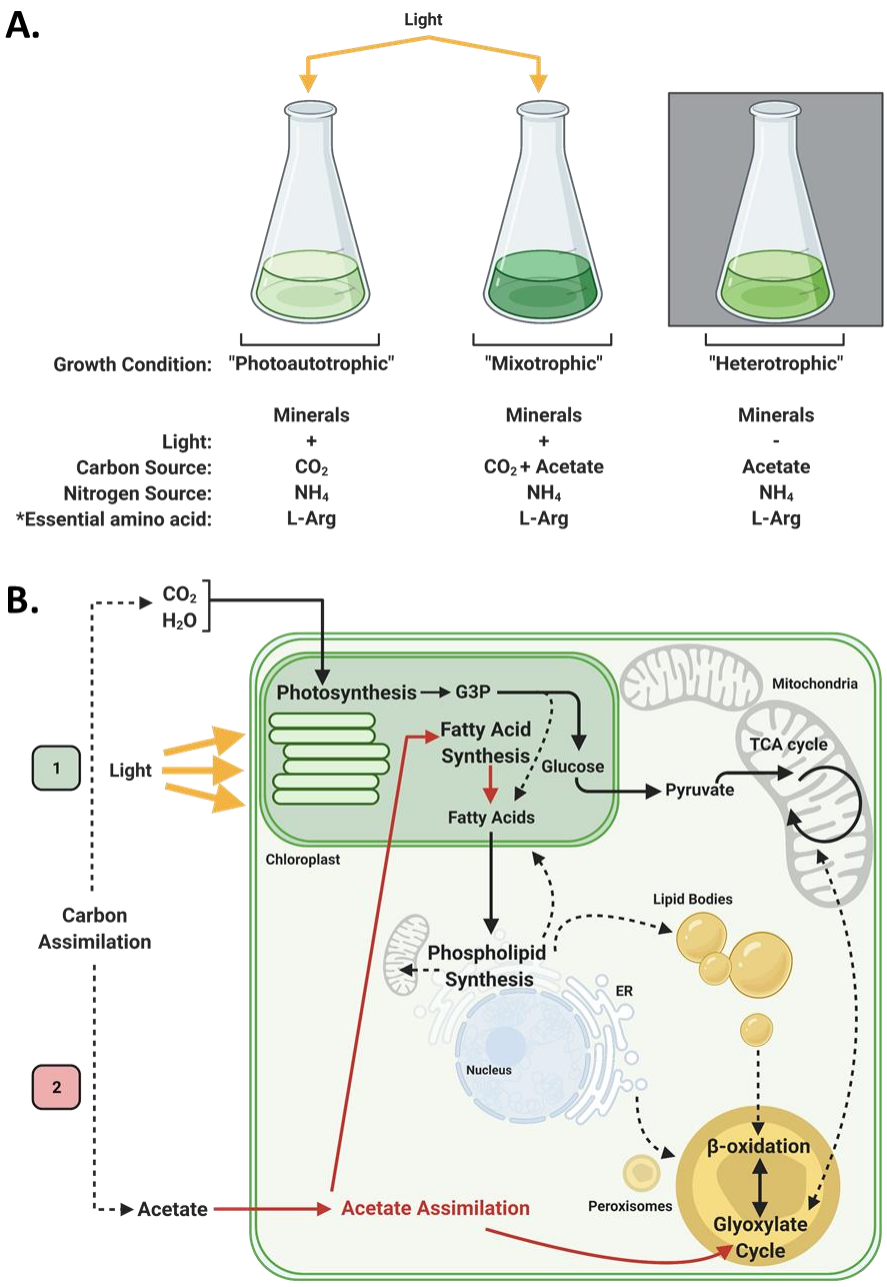


Figure 28 Limited flexibility of *Chlamydomonas* carbon sources.

A) *Chlamydomonas* culture requirements simplified for different growth conditions. Generally, *Chlamydomonas* can be cultured in the light lacking a supplemental carbon source and therefore using only CO_2 (photoautotrophic). Mixotrophic growth occurs when CO_2 is fixed, and an organic carbon source (acetate) is utilized. Heterotrophic or "respiratory" conditions occur when growth can only be due to assimilation of an organic carbon source. B) For sustained growth, *Chlamydomonas* can utilize either CO_2 (in the

presence of light) and/or acetate as supplemental, organic carbon source. When CO_2 is fixed through the Calvin-Benson-Bassham cycle, the cell can produce all carbon containing molecules necessary to support growth of *Chlamydomonas*. Respiratory growth requires function of the glyoxylate cycle for acetate assimilation but acetate, once converted to acetyl-CoA is used in many biochemical reactions. Diagrams created with BioRender.com.

The popularity of *Chlamydomonas* has been fueled by the fact that it naturally produces triacyclglycerols (TAGs), an energy rich carbon storage compound. The mechanisms controlling TAGs accumulation and mobilization has attracted considerable attention in the field of bioenergy as TAGs are biodiesel precursors and thus are useful as fossil fuel substitutes [193, 194]. In *Chlamydomonas*, fatty acid synthesis occurs in the plastid and although the debate is ongoing pertaining to the specific mechanism of transport, fatty acids are subsequently trafficked to the ER for modification (i.e. phospholipid synthesis) and/or redistribution to other cellular compartments. Excess fatty acids are stored as TAGs which accumulate in the cytoplasm as lipid bodies [192, 195, 196]. In instances of high energy demand (such as absence of light for photosynthesis), TAGs can be transported to peroxisomes for catabolism [197, 198]. This process is performed by fatty acid oxidation (also called β-oxidation) which occurs exclusively in the peroxisome in plants, including *Chlamydomonas*. Peroxisomes are a ubiquitously present but functionally diverse class of organelles [199]. In yeast and mammalian model systems, peroxisome biogenesis, structure, and function have been extensively studied but only a few investigations of peroxisome biology in green algae have been undertaken. Since peroxisomes are the sole site of fatty acid oxidation, mutants deficient in peroxisome function would be expected to have elevated lipids. However, whereas a plethora of mutants have been isolated manifesting increased TAG accumulation, those generally affecting peroxisome function have been conspicuously absent. To date, only a few mutants affecting β-oxidation in *Chlamydomonas* having been described [19].

We previously conducted a forward genetic screen for CI assembly mutants based on the slow growth in the dark (SID) phenotype common to *Chlamydomonas* CI mutants. Through this work, the *amc12(#6E9)* mutant was isolated manifesting a decrease in CI assembly and activity [72] (as well as decrease in soluble arm subunits, unpublished). Accompanying the CI assembly defect in *amc12* was a pleiotropic decrease in respiratory complex activities (CII+CIII, and CIV) [72]. Intriguingly, mitochondrial encoded transcripts of CI, CIII, and CIV subunits were also decreased (unpublished). However, a second, independently segregating mutation was subsequently found unlinked to the insertional cassette clouding interpretation of these results (unpublished).

Here we report the identification and characterization of the single mutant, referred to as *lcla1*, isolated from the original *amc12* double mutant (*lcla1*, *sid2*). Like the original mutant strain, a strain carrying the single *lcla1* mutation, also was affected for growth in respiratory conditions (dark plus acetate as a carbon source). However, unlike *amc12* it also grew more slowly in mixotrophic conditions (light plus acetate) and this phenotype was further modulated by mutations in the *ARG7* gene, a commonly used nutritional marker. Interestingly, mixotrophic growth of the *Chlamydomonas* isocitrate lyase (*icl*) mutant was also modulated by mutations inactivating ARG7. As in the *icl* mutant, *lcla1* had increased lipid content and was sensitive to high light [17], a stress condition that induces lipid synthesis in the plastid. Notably, it is also resistant to the fatty acid synthesis inhibitor Cerulenin in high light, suggesting that the light-sensitivity was caused by lipotoxicity. The increased lipid content, inferred by increase in TAGs and cytoplasmic lipid droplets is due to a defect in lipid utilization as the single mutant is unable to use its

lipid reserves when transferred to media lacking an alternative carbon source. Furthermore, supplementation with oleic acid, a long chain fatty acid, is unable to support growth when present as the sole carbon source suggesting that this defect is specific to lipids instead of other forms of stored carbon. Intriguingly, inability to grow on oleic acid is characteristic of peroxisome deficiency in yeast [200].

The insertional cassette was mapped to *Cre07.g329861*, a previously uncharacterized gene encoding a hypothetical protein of unknown function and reintroduction of the WT copy rescued the growth and TAG/lipid body phenotypes. Further sequence analysis identified a putative membrane peroxisome targeting signal (PTS) at the N-terminus and a chimeric fusion protein containing this motif is sufficient to direct an artificial reporter to the ER, the site of *de novo* peroxisome biogenesis (in mixotrophic conditions). Shifting cultures to respiratory conditions, the chimeric reporter redistributes, similar to that of a peroxisome-targeted, fluorescent marker. Here we implicate a novel protein, LCLA1, for Low-Complexity in Lipid Accumulation in algal lipid metabolism, possibly associated with peroxisome biogenesis. While the precise mechanism of action remains to be elucidated, potential functions of LCLA1 are discussed.

4.3 Materials and methods

4.3.1 Strains, media, and culture conditions

The background strain used for insertional mutagenesis was 4C− (*mt⁻; arg7-8*) (Dr. Rochaix, University of Geneva, deposited in Chlamy Resource Center as CC-5590). The *amc12#6E9* mutant (*mt⁻; arg7-8; lcla1::APHVII; sid2*) was obtained from this screen and backcrossed to the wild-type strains CC-4533 (also known as CMJ030, *mt⁻; cw15*) or its isogenic mating partner, CC-5155 four times [11]. Eight wild-type meiotic progeny were retained from the final backcross: 4 [arg+] and 4 [arg-] haploid progeny (WT; *arg7-8*). Likewise, eight *lcla1* mutant haploid progeny were retained: 4 [arg+] (*lcla1::APHVII*) and 4 [arg-] haploid progeny (*lcla1::APHVII; arg7-8*). One of the resulting haploid progeny, *lcla1#1* was chosen as a recipient strain for complementation (described further below). CC-4533 and CC-5155 were chosen as recipient strains for expression of the chimeric fluorescent reporter or controls (described below). CSI_FC2A03 (CC-4533 expressing MDH4-CrVENUS-3xFLAG) and CSI_FC2C02 (CC-4533 expressing Cre16.g683600-CrVENUS-3xFLAG) were obtained from the *Chlamydomonas* Resource Center and used as mitochondrial and cytosolic markers, respectively, for confocal fluorescence microscopy [201]. All strains used in this study are summarized in full in table 3.

Strain Name	Alias	Genotype	Description
WT	4C-	*mt⁻; arg7-8*	WT strain used for insertional mutagenesis yielding *amc8 - amc13* [72], CC-5590
amc12	*amc12#6E9*	*mt⁻; arg7-8; lcla1::APHVII; sid2*	*amc12#6E9* original double mutant, SID [72], CC-5611
WT	137c+	*mt⁺*	Obtained from *Chlamydomonas* Resource Center, is wild type background of *icl1* mutant, yellow in the dark [17], CC-125
icl	*icl1#919*	*mt⁺; icl::APHVII*	Generously provided by Dr. Claire Remacle, dies in respiratory conditions (*dk*-) [17]
icl(ICL)	*icl1(ICL1)#921*	*mt⁺; icl::APHVII; ICL*	Generously provided by Dr. Claire Remacle [17]
icl(ICL-FLAG)	icl1(ICL1-FLAG)#937	*mt⁺; icl::APHVII; ICL-FLAG::APHVIII*	Generously provided by Dr. Claire Remacle [15]
***icl*, [arg+] spore**	*icl1#PH45-230*	*icl::APHVII*	Obtained by cross of *icl(ICL-FLAG)* x WT, [arg-] spore, dies in respiratory conditions (*dk*-)
***icl*, [arg+] spore**	*icl1#PH45-243*	*icl::APHVII*	Obtained by cross of *icl(ICL-FLAG)* x WT, [arg-] spore, dies in respiratory conditions (*dk*-)
***icl*, [arg-] spore**	*icl1#PH43-143*	*icl::APHVII; arg7-8*	Obtained by cross of *icl(ICL-FLAG)* x WT, [arg-] spore, dies in respiratory conditions (*dk*-)
***icl*, [arg-] spore**	*icl1#PH45-250*	*icl::APHVII; arg7-8*	Obtained by cross of *icl(ICL-FLAG)* x WT, [arg-] spore, dies in respiratory conditions (*dk*-)
***icl*, [arg-] spore**	*icl1#PH45-235*	*icl::APHVII; arg7-8*	Obtained by cross of *icl(ICL-FLAG)* x WT, [arg-] spore, dies in respiratory conditions (*dk*-)

Continued

Table 3 List of strains used in this study (continued table).

Table 3 continued

Strain Name	Alias	Genotype	Description
WT	CC-4533	*mt⁻; cw15*	WT used for phenotypic comparison (also called CMJ030) [11, 202]
WT	CC-5155	*mt⁺; cw15*	WT isogenic mating partner of CC-4533
WT, [arg-] spore	WT#PH25-1	*mt⁻; arg7-8*	WT, [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg-] spore	WT#PH25-2	*mt⁺; arg7-8*	WT, [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg-] spore	WT#PH25-3	*arg7-8*	WT, [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg-] spore	WT#PH25-4	*mt⁻; arg7-8*	WT, [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg+] spore	WT#PH25-5	-	WT, [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg+] spore	WT#PH25-6	-	WT, [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg+] spore	WT#PH25-7	-	WT, [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*
WT, [arg+] spore	WT#PH25-8	-	WT, [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155, behaves *cwd*

Table 3 continued

Strain Name	Alias	Genotype	Description
***amc9*, [arg+] spore**	*amc9#PH40-31*	*mt^+; nuo5::APHVII*	*amc9*, [arg+] spore resulting obtained from 4th backcross of *amc9* (*nuo5-null*) to CC-4533/CC-5155 [72]
***amc9*, [arg+] spore**	*amc9#PH42-48*	*mt^+; nuo5::APHVII*	*amc9*, [arg+] spore resulting obtained from 4th backcross of *amc9* (*nuo5-null*) to CC-4533/CC-5155 [72]
***amc9*, [arg+] spore**	*amc9#PH41-18*	*mt^+; nuo5::APHVII*	*amc9*, [arg+] spore resulting obtained from 4th backcross of *amc9* (*nuo5-null*) to CC-4533/CC-5155 [72]
***amc9*, [arg+] spore**	*amc9#PH39-3*	*mt^-; nuo5::APHVII*	*amc9*, [arg+] spore resulting obtained from 4th backcross of *amc9* (*nuo5-null*) to CC-4533/CC-5155 [72]
***amc9*, [arg-] spore**	*amc9#PH42-16*	*mt^-; arg7-8; nuo5::APHVII*	*amc9*, [arg-] spore resulting obtained from 4th backcross of *amc9* (nuo5-null) to CC-4533/CC-5155 [72]
***amc9*, [arg-] spore**	*amc9#PH39-26*	*mt^+; arg7-8; nuo5::APHVII*	*amc9*, [arg-] spore resulting obtained from 4th backcross of *amc9* (nuo5-null) to CC-4533/CC-5155 [72]
***amc9*, [arg-] spore**	*amc9#PH40-11*	*mt^+; arg7-8; nuo5::APHVII*	*amc9*, [arg-] spore resulting obtained from 4th backcross of *amc9* (nuo5-null) to CC-4533/CC-5155 [72]
***amc9*, [arg-] spore**	*amc9#PH41-27*	*mt^-; arg7-8; nuo5::APHVII*	*amc9*, [arg-] spore resulting obtained from 4th backcross of *amc9* (nuo5-null) to CC-4533/CC-5155 [72]

Table 3 continued

Strain Name	Alias	Genotype	Description
***lcla1*, [arg-] spore**	*lcla1#PH25-1*	*mt⁻; arg7-8; lcla1::APHVII*	*lcla1* [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155; recipient strain for WT copy *LCLA1* and controls, is now suppressed
***lcla1*, [arg-] spore**	*lcla1#PH25-2*	*mt⁻; arg7-8; lcla1::APHVII*	*lcla1* [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155
***lcla1*, [arg-] spore**	*lcla1#PH25-3*	*mt⁻; arg7-8; lcla1::APHVII*	*lcla1* [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155
***lcla1*, [arg-] spore**	*lcla1#PH25-4*	*mt⁻; arg7-8; lcla1::APHVII*	*lcla1* [arg-] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155; alternative *lcla1#PH25-1*
***lcla1*, [arg+] spore**	*lcla1#PH25-5*	*mt⁻; lcla1::APHVII*	*lcla1* [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155; used initaially, is now suppressed
***lcla1*, [arg+] spore**	*lcla1#PH25-6*	*mt⁺; lcla1::APHVII*	*lcla1* [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155; alternative *lcla1#PH25-5*
***lcla1*, [arg+] spore**	*lcla1#PH25-7*	*mt⁻; lcla1::APHVII*	*lcla1* [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155
***lcla1*, [arg+] spore**	*lcla1#PH25-8*	*mt⁻; lcla1::APHVII*	*lcla1* [arg+] spore resulting obtained from 4th backcross of *amc12#6E9* to CC-4533/CC-5155
lcla1(EV)	*lcla1#PH25-1 (pCB412)#1*	*mt⁻; lcla1::APHVII; ARG7*	Empty vector transformant (pCB412) but *ARG7* now partially suppressed
lcla1(EV)	*lcla1#PH25-1 (pCB412)#2*	*mt⁻; lcla1::APHVII; ARG7*	Initially used as EV (pCB412) transformed control, SID phenotype now partially suppressed
lcla1(EV)	*lcla1#PH25-1 (pCB412)#3*	*mt⁻; lcla1::APHVII; ARG7*	Alternate empty vector (pCB412) transformed control

Table 3 continued

Strain Name	Alias	Genotype	Description
lclal1(64B9)	*lcla1#PH25-1 (64B9<ScaI)#28*	*mt⁻, lcla1::APHVII; LCLA1::ARG7*	Linearized with *Sca*I prior to transformation (restricts *bla* gene)
lcla1(64B9<Bar**I**)	*lcla1#PH25-1 (64B9<SgrDI)#*	*mt⁻, lcla1::APHVII; LCLA1::ARG7*	Linearized with *Bar*I prior to transformation (restricts *Cre07.g329882*)
lcla1(64B9<Sgr**DI**)	*lcla1#PH25-1 (64B9<BarI)#*	*mt⁻, lcla1::APHVII; ARG7*	Linearized with *Bar*I prior to transformation (restricts *LCLA1*)
lcla1(LCLA1)	*lcla1#PH25-1 (pADC2)#19*	*mt⁻, lcla1::APHVII; LCLA1::ARG7*	*lcla1* complemented strain
lcla1(LCLA1)	*lcla1#PH25-1 (pADC2)#20*	*mt⁻, lcla1::APHVII; LCLA1::ARG7*	*lcla1* complemented strain
lcla1(LCLA1)	*lcla1#PH25-1 (pADC2)#21*	*mt⁻, lcla1::APHVII; LCLA1::ARG7*	*lcla1* complemented strain
lcla1(LCLA1-3xFLAG)	*lcla1#PH25-1 (pADC10)#20*	*mt⁻, lcla1::APHVII; LCLA1-3xFLAG::ARG7*	Complemented, for detection of *LCLA1-3xFLAG*
lcla1(LCLA1-3xFLAG)	*lcla1#PH25-1 (pADC10)#25*	*mt⁻, lcla1::APHVII; LCLA1-3xFLAG::ARG7*	Complemented, for detection of *LCLA1-3xFLAG*
lcla1(LCLA1-6xFLAG)	*lcla1#PH25-1 (pADC11)#17*	*mt⁻, lcla1::APHVII; LCLA1-6xFLAG::ARG7*	Complemented, for detection of *LCLA1-6xFLAG*
lcla1(LCLA1-6xFLAG)	*lcla1#PH25-1 (pADC11)#42*	*mt⁻, lcla1::APHVII; LCLA1-6xFLAG::ARG7*	Complemented, for detection of *LCLA1-6xFLAG*

Table 3 continued

Strain Name	Alias	Genotype	Description
Peroxisome-CrVENUS	*uvm4*(pOpt_CIS2-mVenus_paro)#7	*mt*$^{+}$, *cw-, CIS2-N25-mVENUS::APHVIII*	*Uvm4* strain is more sensitive to antibiotics, only grows ≤ 5 µg/ml HygB [15]
Mito-CrVENUS	CSI_FC2A03	*mt*$^{-}$*; MDH4-CrVENUS::APHVIII*	CC-4533(pLM005-MDH4-CrVENUS-3xFLAG)
Cyto-CrVENUS	CSI_FC2C02	*mt*$^{-}$*, Cre16.g683600-CrVENUS::APHVIII*	CC-4533(pLM005-Cre16.g683600- CrVENUS-3xFLAG)
***lcla1*, Mito-CrVENUS**	*lcla1#PH25-6* x FC2AO3 #6B	*lcla1::APHVII; MDH4-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2A03 (Mito-CrVENUS)
***lcla1*, Mito-CrVENUS**	*lcla1#PH25-6* x FC2AO3 #5D	*lcla1::APHVII; MDH4-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2A03 (Mito-CrVENUS)
***lcla1*, Mito-CrVENUS**	*lcla1#PH25-6* x FC2AO3 #8D	*lcla1::APHVII; MDH4-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2A03 (Mito-CrVENUS)
***lcla1*, Cyto-CrVENUS**	*lcla1#PH25-6* x FC2CO2 #5B	*lcla1::APHVII; Cre16.g683600-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2C02 (Cyto-CrVENUS)
***lcla1*, Cyto-CrVENUS**	*lcla1#PH25-6* x FC2CO2 #5E	*lcla1::APHVII; Cre16.g683600-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2C02 (Cyto-CrVENUS)
***lcla1*, Cyto-CrVENUS**	*lcla1#PH25-6* x FC2CO2 #8A	*lcla1::APHVII; Cre16.g683600-CrVENUS::APHVIII*	Spore resulting from cross between *lcla1#6* and FC2C02 (Cyto-CrVENUS)

Table 3 continued

Strain Name	Alias	Genotype	Description
Cyto-CrVENUS	CC-4533(pMO449) #I-A10	*mt⁻, CrVENUS::APHVIII*	Independent CC-4533(pMO449) transformant, cytoplasmic localization of CrVENUS marker
Cyto-CrVENUS	CC-4533(pMO449) #I-B10	*mt⁻, CrVENUS::APHVIII*	Independent CC-4533(pMO449) transformant, cytoplasmic localization of CrVENUS marker
Cyto-CrVENUS	CC-4533(pMO449) #I-C10	*mt⁻, CrVENUS::APHVIII*	Independent CC-4533(pMO449) transformant, cytoplasmic localization of CrVENUS marker
LCLA1^{S}-N72-CrVENUS	CC-4533(pADC9) #II-G4	*mt⁻, LCLA1^{S}-N72 CrVENUS::APHVIII*	Independent CC-4533(pADC9) transformant, PCR-positive, chimeric protein expression confirmed via immunoblot
LCLA1^{S}-N72-CrVENUS	CC-4533(pADC9) #II-F7	*mt⁻, LCLA1^{S}-N72 CrVENUS::APHVIII*	Independent CC-4533(pADC9) transformant, PCR-positive, chimeric protein expression confirmed via immunoblot
LCLA1^{S}-N72-CrVENUS	CC-4533(pADC9) #II-F1	*mt⁻, LCLA1^{S}-N72 CrVENUS::APHVIII*	Independent CC-4533(pADC9) transformant, PCR-positive, chimeric protein expression confirmed via immunoblot

Strains were maintained monthly in low light (2-5 μmol photons m^{-2} s^{-1}) on selective media, either TAP (Tris Acetate Phosphate) or on TARG (TAP+Arg) media for arginine auxotrophic strains. TAP media is composed of 20 mM Tris base, 17 mM acetic acid, 113 mM phosphate supplemented with trace minerals as previously described [1, 8, 203]. Where indicated, L-arginine was provided at a final concentration of 400 μg/ml (1.9 mM). For strains expressing antibiotic resistance markers, either Paromomycin sulfate (Pm, RPI, #P11000-25.0) and/or Hygromycin B (HygB, Invitrogen, #10687010) were each used at a final concentration of 25 μg/ml. Unless otherwise indicated, cultures were grown with constant illumination at 25-35 μmol photons m^{-2} s^{-1} (i.e. without diurnal light/dark cycles) except in respiratory conditions. In accordance with our laboratory conditions, we define normal light conditions as 30-50 μmol m^{-2} s^{-1} and low light conditions correspond to 2-5 μmol m^{-2} s^{-1}. Solid medium contains 1.5% (w/v) select agar (Invitrogen, #30391049). Liquid cultures were performed where indicated with constant light (40-60 μmol photons m^{-2} s^{-1}) and shaking (180 rpm). All cultures were maintained at 25ºC.

Chemically competent *Escherichia coli* DH5α strains were used for molecular cloning (ThermoFisher Scientific, #C404003). *E. coli* was grown at 37°C in Luria-Bertani (LB) broth pH 7.2 and bacteriological agar [69].

4.3.2 Genetic analysis

SID mutants were backcrossed as described in [1] selecting for Hyg^R imparted by the insertional cassette. Mating type was determined by genotyping PCR with previously determined methods [8, 72, 74, 132] and vegetative cells of opposite mating type refreshed on selective media prior to crossing. To induce gametogenesis, vegetative cells from solid media were resuspended in TAP liquid medium lacking nitrogen (TAP-N). These preparations were cultured in the dark at 25°C and 180 rpm for 4-6 hours. Gametes of opposite mating type were then mixed (1:1) in a small petri dish and left overnight at room temperature and constant ambient light. To allow zygote maturation, the overnight mixture of mated gametes was transferred to solid TAP-N media containing 3% agar and incubated for 5-7 days (constant light ~15 μmol photons m^{-2} s^{-1}, 25°C). Following incubation, the unmated gametes were scraped away with a dull, sterile blade leaving behind the mature zygotes embedded in the agar. Zygotes were then collected and transferred to nitrogen replete media (TARG) to permit germination in low light. Individual colonies were subcloned 2-3x to ensure clonality of selected spores. This was repeated four rounds before performing tetrad analysis as described previously [204].

4.3.3 Genotyping PCR

Diagnostic PCRs were performed using GoTaq polymerase (Promega, #M3008) supplemented with 2.5% (v/v) with DMSO but otherwise in accordance with manufacturer's instructions. Forward and reverse primers were added each to a final concentration of 1 μM (see table 4 for all primer sequences). Since the *Chlamydomonas* nuclear genome is GC-rich, a denaturation step of 98°C was performed instead of 95°C.

Annealing temperature varied depending on primer T_m and 35-40 cycles were selected for amplification. PCR products were resolved on 1-2% agarose gels depending on amplicon size in 0.5x TBE buffer. Ethidium bromide was added to agarose gels (0.5-1.0 µl from 10 mg/ml stock per 100 ml) for UV detection of ethidium bromide-intercalated PCR products.

Table 4 List of primers used in this study.

Primer Name	Sequence (5' to 3')	Target
PH3	GGAGTCCCGGGATGGATTAAGG	64B9 5'border genomic segment insert in pCB412
64B9-480R	TCGTGCTTCCGACATTTACC	
64B9-35317F	GTGGAGTAAAGCGTTAGGAG	64B9 3'border genomic segment insert in pCB412
pBR329-3	GACAGCATCGCCAGTCACTATGG	
PH1	GTTTTCGTTCCACTGAGCGTCAG	pCB412 backbone
PH2	ATGCCCGCGTTGCAGGCCATGC	
6E9-RT2-F1	CAAGAGCAGCCCTGCGTGTG	5' border iHyg3 insertional cassette in *Cre07.g329861*
APH7-R5	CGGTCGAGAAGTAACAGGG	
APH7-F8	ACTGCTCGCCTTCACCTTC	3' border iHyg3 insertional cassette in *Cre07.g329861*
Cre07.g329861-exon1-R2	CAACTGGTCTGCGCTGAG	
Cre07.g329861-8832F	TGTACAACCGCATGTACGTG	*Cre07.g329861*
Cre07.g329861-9384R	GCAGCATCAGGCAATATAGC	
Cre07.g329861-9401F	CAAGGCTCCGGCAACGTAAG	Flank *Sgr*DI restriction site used for insertion of fragment encoding 3xFLAG
Cre07.g329861-9762R	TGCCACACGCGAACTAGCTG	
NUO5-E2L	ACTACAAGCAGAGCGCCATC	*NUO5*
NUO5-E3R	CATCTCGCCCAGTGTGAAC	
Mid-up	ATGGGCCTGTTTCTTAGC	Mid mating type locus for mt^-
Mid-low	CTACATGTGTTTCTTGACG	
Fus-up	ATGCCTATCTTTCTCATTCT	Fus mating type locus for mt^+
Fus-low	GCAAAATACACGTCTGGAAG	
LCLA1-S-N72-F	ATGCTGCGGGCGGAGG	Amplified 5' end of LCLA1 encoding 1st 72 amino acids of LCLA1^S
LCLA1-S-N72-R	CGCTCCCGCCTCTTGAGGTA	
APH8-F5	GTGTTCCGCGGCGTTCCTG	*APHVIII* in pMO449 and derivative vectors
pMO520-R9	ATACGCAAACCGCCTCTCCC	
pMO520-F7	GCAGCGTGCTTGCAGATTTG	Flank *Hpa*I cloning site in pMO449 and derivative vectors
pLM005-genotyping R1	AACTTGTGGCCGTTCACGTC	

4.3.4 Glass bead transformation and genetic complementation

An indexed cosmid-based genomic library containing regions covering the entirety of the *Chlamydomonas* nuclear genome was provided by Dr. Jean David Rochaix's laboratory [76, 134]. The 64B9 cosmid was found to contain the *LCLA1* gene (*Cre07.g329861*) by diagnostic PCR and the size of the *LCLA1* genomic region in 64B9 was identified by mapping the sequenced borders of the insert to the *Chlamydomonas* genome (Augustus v5.5, Figure 29A). Cosmids were purified using the NucleoBond Xtra Midi kit for transfection-grade plasmid DNA (Macherey Nagel, # 740410.50) according to manufacturer's protocol. Manipulation of transforming DNA was designed using CloneManager and SnapGene programs. The 64B9 cosmid was further manipulated by restriction digest and different versions introduced into the backcrossed *lcla1* (Figure 30 and 31). 64B9 was linearized with the *Sca*I restriction enzyme (ThermoFisher, #FD0434) for 1-2 hrs at 37^{o}C prior to transformation. Restricted versions of the 64B9 cosmid in which either the *LCLA1* gene or downstream gene (*Cre07.g329882*) were disrupted by digest with *Sgr*DI (ThermoFisher, #ER2031 - digest 2 hrs at 37^{o}C) or *Bar*I (Sibenzyme, #E547 - digest 2 hrs at 37^{o}C) respectively and used directly for transformation of the *lcla1* mutant. Digest of 64B9 with *Bst*BI and relegation removed 20.5 kb of 64B9 including the entire downstream gene 281 bp after the predicted translation start site (i.e. 93.67 of 1354 codons in *Cre07.g329882*, Figure 32) and herafter named pADC2. pCB412, the parent vector used in the development of the cosmid library, was also linearized by digestion with *Sca*I and transformed into the *lcla1* mutant as an empty vector control. Transformants were

selected for arginine prototrophy after 1-2 weeks on TAP+HygB plates and subcloned 2-3x to ensure purity for further experimentation. All primers are included in table 4 and transforming DNA summarized in table 5.

Table 5 List of vectors used in this work.

Vector	Alias	Selectable marker	Comment
pCB412	EV	*ARG7*	Empty vector used in creation of cosmid library [76, 205]
64B9	-	*ARG7*	pCB412 containing the genomic segment including the full genes: *Cre07.g329861* and *Cre07.g329882*
pADC2	64B9<*Bst*BI	*ARG7*	64B9 further restricted to only *Cre07.g329861* with digestion of *Bst*BI and re-ligation
pADC10	pADC2.1-3xFLAG#27	*ARG7*	3xFLAG inserted into unique *Sgr*DI site of pADC2 (in frame)
pADC11	pADC2.1-6xFLAG#22	*ARG7*	6xFLAG inserted into unique *Sgr*DI site of pADC2 (in frame)
pMO449	Cyto-CrVENUS-3xFLAG	*APHVIII*	*Chlamydomonas* expression vector in which the fluorophore is transcriptionally linked to Pm^R marker [206]
pADC8	$LCLA1^L$-N114-CrVENUS-3xFLAG	*APHVIII*	Sequence of the start codon in LCLA1-Long containing the first 114 amino acids cloned via *Hpa*I site in pMO449
pADC9	$LCLA1^S$-N72-CrVENUS-3xFLAG	*APHVIII*	Sequence of the start codon in LCLA1-Short containing the first 72 amino acids cloned via *Hpa*I site in pMO449

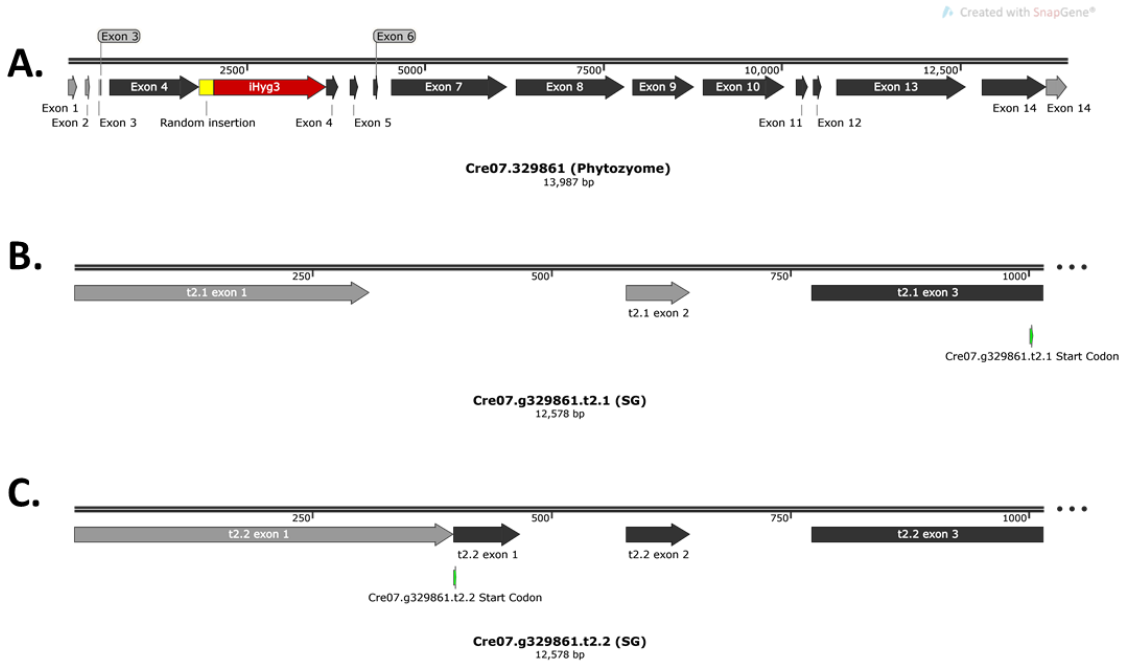


Figure 29 Insertional cassette maps to novel gene.

A) Gene model according to the latest *Chlamydomonas* genome version 5.5 available through the Phytozome web browser (https://phytozome.jgi.doe.gov/pz/portal.html). Insertion of the hygromycin resistance cassette marked in red (iHyg3). Accompanying carrier DNA inserted upstream of along with iHyg3 cassette colored yellow. B) Updated gene model for 5' shows alternative splice of 1st exon resulting in shorter transcript (Cre07.g329861.t2.1 encoding LCLA1^{S}) skipping over intended start codon. Second start codon in exon 3 (formerly exon 4) of *Cre07.g329861*. C) Updated gene model for Cre07.g329861.t2.2 (encoding LCLA1^{L}). Scale marked above each exon; introns not shown. Models made with SnapGene Viewer software.

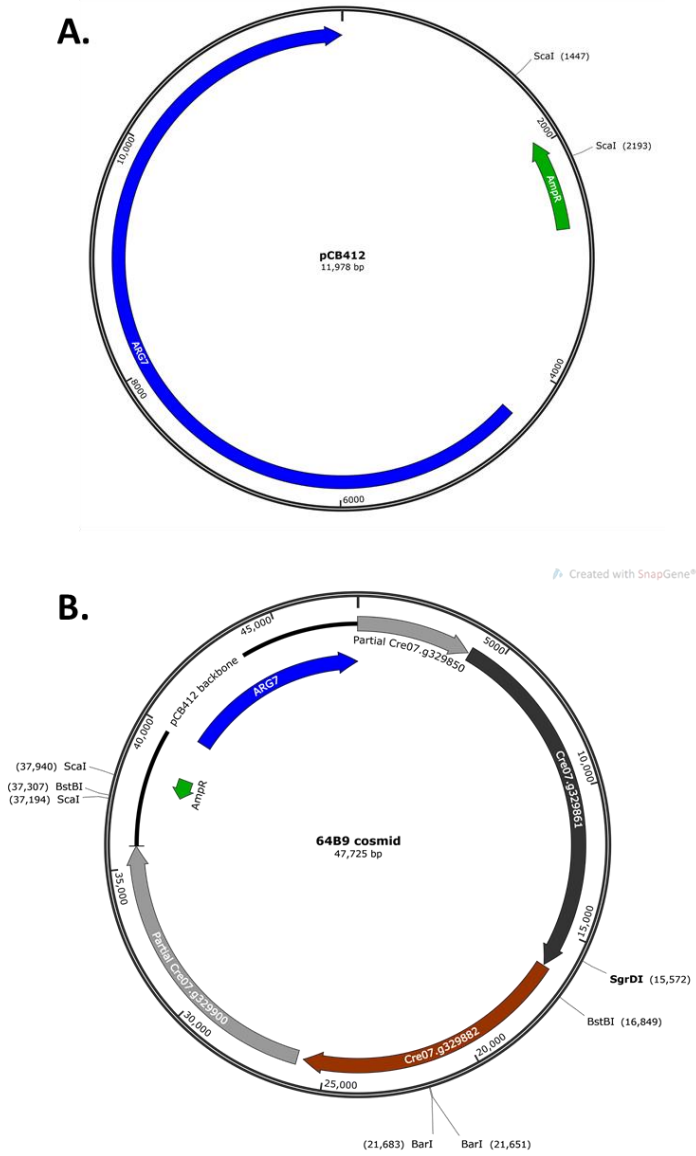


Figure 30 Cosmid vector maps used in complementation.

A) pCB412 parent vector used to generate the *Chlamydomonas* cosmid library. B) Cosmid 64B9 carrying a genomic segment of chromosome 7 containing the full *Cre07.g329861* gene disrupted in the *lcla1* mutant. Restriction sites used in split gene complementation test or in cloning of pADC2 are indicated on vector map. *Sgr*DI site is also used for insertion of sequence encoding 3xFLAG epitope tag. Vector maps (not to scale) made using SnapGene Viewer.

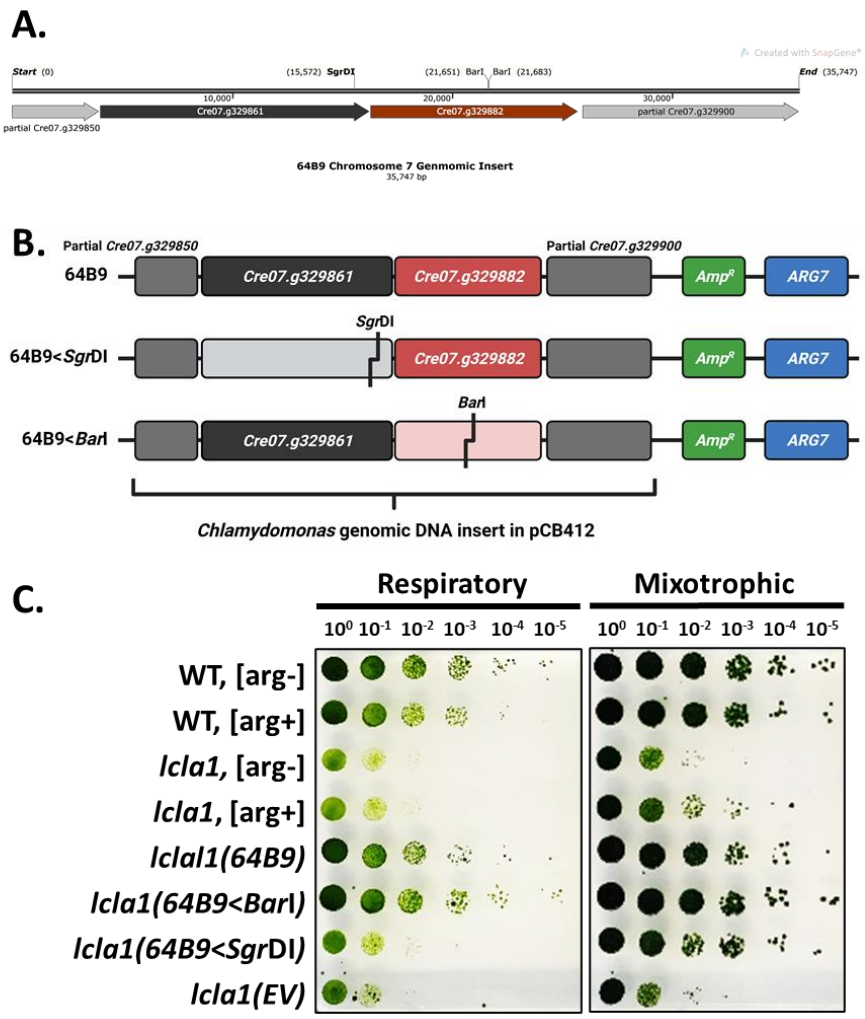


Figure 31 Split gene complementation test confirms *Cre07.g329861* disruption cause of SID phenotype.

A) 64B9 genomic segment map with select restriction sites indicated. B) Model for split gene complementation test. C) Tenfold dilution series of representative transformants from (B) on TARG media incubated in the light (30 μmol photons m^{-2} s^{-1}) or in the dark for 10 days. WT, [arg-] was obtained from final backcross yielding *lcla1*. WT, [arg+] is CC-4533. Transformed strains (with parentheses) are all [arg+] due to *ARG7* selectable marker of cosmid (see Table 5). Map in (A) made using SnapGene Viewer and model for (B) created using BioRender.com.

The glass bead method [76, 207] was used for introduction of foreign DNA with several noteworthy modifications. The strains to be transformed were refreshed 1-2x on selective media then a small loop-full used to inoculate Erlenmeyer flasks containing 250 ml of TAP media (supplemented with 40 mM sorbitol for *cw$_d$* strains and/or 400 µg/ml Arg as needed) and incubated for 3-5 days until reaching mid exponential phase (2-6x10^6 cells/ml). Cells were then collected by centrifugation at 1,500 x g for 5 mins at room temperature and the pellet resuspended to a final cell density of 1x10^8 cells/ml in a gamete autolysin preparation for cell wall digestion. Complete cell wall digestion (45 min – 2 hr depending on strain) was monitored intermittently by spotting of 2.5 µl cell suspension on glass slides with 2.5 µl 0.1% triton x 100 (Fisher, #BP151-500) and coverslip. Lysis of detergent-treated cells (observed at 100x total magnification using a compound light microscope) indicated sufficient cell wall digestion.

Following autolysin treatment, the cell suspension was centrifuged again, autolysin discarded, and cell pellet resuspended in TAP+arg+Sorb (40 mM sorbitol) to a final cell density of 1x10^8 cells/ml. The transformation reaction was then set up as follows in sterile 1.5 ml microcentrifuge tubes containing 300 µg sterile glass beads: in order, (1) 100 µl 20% polyethylene glycol (PEG 8000), (2) 300 µl cells at 1x10^8 cells/ml, (3) 2.5 µl filter-sterilized sheared herring sperm DNA (10mg/ml), and (4) transforming DNA (1-5 ug linearized by restriction digest). The transformation reactions were then vortexed at max speed for 30 seconds and approximately 300 µl of supernatant transferred immediately (without glass beads). Recovery incubations were performed for 24-48 hours in 30 ml TARG+Sorb (in 125 ml Erlenmeyer flasks) shaking at 180 rpm, 25^{o}C, with constant

illumination at 30-40 µmol photons. After recovery incubation, transformant cultures were pelleted by centrifugation at 1,500 x *g* for 5 mins at room temperature and resuspended in 1 ml TARG+Sorb. If arginine prototrophy was used for selection, the cell pellet was washed 3x with 5 ml TAP media (lacking arginine) or sterile water prior to plating. 500 µl of cell suspension was then plated by glass spreader onto selective media allowed to dry in sterile hood. If successful, transformant colonies were visible within 5-10 days incubation in normal growth conditions (any colonies that appeared after two weeks were often not true transformants). Transformant colonies were picked and subcloned 2-3 times on selective media to ensure purity before confirmation of the presence of the transforming DNA by diagnostic PCR. Due to positional effects of insertion, multiple independent transformants were evaluated.

4.3.5 Growth assessment

Growth in different nutrient or light conditions was assessed by ten-fold dilution series as described previously [8, 72, 132]. Strains to be tested were refreshed 1-2x on selective media then resuspended in 500 µl sterile media (or ddH_2O). Samples were normalized to a starting optical density at 750 nm (OD_{750}) of 5.0 or 2.0 and diluted 1:10 sequentially in 96-well plates for a total of 5 serial dilutions (plus undiluted). Five µl from each dilution were spotted onto solid media and incubated in the indicated light conditions. Plates were imaged periodically in 1-week or 5-day intervals as indicated. For assessing photoautotrophic growth, dilutions were performed in minimal media (lacking acetate as a

carbon source) and spotted onto TP+arg media in the same manner. Ten-fold dilution series were often performed in 3 light conditions: dark (0.0 μmol photons m^{-2} s^{-1}), normal light (30-50 μmol photons m^{-2} s^{-1}), and high light (200-300 μmol photons m^{-2} s^{-1}).

To investigate whether acetate, ammonium, or arginine were limiting for growth, concentrations of each were incrementally varied as indicated while controlling for pH (7.0). Likewise, when testing for sensitivity or resistance to different drugs such as canavanine (Sigma, #C1625-250MG) or cerulenin (Sigma, #C2389-5MG), each were added to the indicated final concentrations. Drugs were either purchased as sterile preparations or a stock solution was filter-sterilized (0.2 μM pore size) before adding to autoclaved media cooled to below 60°C.

Growth curve analysis was performed as described previously [72, 132]. Briefly, strains were refreshed on selective media for 2-5 days before inoculating starter cultures of 200 ml TARG in 500 ml Erlenmeyer flasks. Starter cultures were maintained in mid-exponential phase by daily dilutions for 3-5 days before counting cell density and collection via centrifugation (5 min at 2,000 *x g*, 25°C). Cultures for growth curve analysis were inoculated to a starting cell density of $1x10^5$ cells/ml (in total volume of 35ml TARG). At approximately 24-hour increments, Optical density (OD_{750}) was measured spectrophotometrically and cell density determined. For each measurement, 2 ml culture was required and replaced by fresh TARG media. All culture conditions were as described above.

4.3.6 Oleic acid feeding

Oleic acid feeding was performed as described previously with several modifications [19]. Briefly, starter cultures of 200 ml were inoculated from 2-5 day old plates and maintained in mid-exponential phase with daily dilutions as for growth curve analysis. Prior to carbon source dropout, cell density was counted via hemocytometer. Cells were pelleted by centrifugation (5 min at 2,000 *x g*, 25ºC) washed with TP media (pH 7.8), then collected again via centrifugation again (5 min at 2,000 *x g*, 25ºC). Samples were resuspended in TP media to starting concentration of $2x10^6$ cells/ml in foil-wrapped flasks. Oleic acid (Alfa Aesar, #AA3199714) was added where indicated to a final concentration of 0.5 mM at inoculation and again 24 hrs later. Cell density was counted via hemocytometer twice daily and plotted as a function of time.

4.3.7 Protein extraction and immunodetection

Protein samples were prepared from mid-exponential phase liquid cultures by freeze thaw method as described previously [208]. Briefly, cells were pelleted by centrifugation (5 min at 2,000 *x g*, 25ºC) and supernatant discarded. Cell pellets were resuspended in 10mM NaPi buffer (pH 7.0) and transferred to 1.5 ml microcentrifuge tubes. Cells were recollected by centrifugation (same conditions) and supernatant removed before storing at -80ºC. For whole-cell extract, pellets were thawed, and protein

concentration measured by Bradford assay. Samples were normalized with 10mM NaPi buffer (pH 7.0), treated with 1,4-dithiothreitol (DTT, to final concentration of 50 mM) and 2-mercaptoethanol (BME, 10% v/v final concentration) and diluted in 2x Laemmli sample buffer (Laemmli, 1970). 2.5 – 20 µg protein sample was loaded per lane as indicated and resolved by SDS-PAGE on 10 or 15% (w/v) polyacrylamide gels as described before [78, 208].

For collection of soluble or membrane fractions, a second freeze/thaw step is required to release cellular compartments as described previously [208]. In this case, high speed centrifugation (5 min at ~16,000 *x g*, 4°C) of samples after the secund thaw is required. The supernatant, containing the soluble fraction is transferred to new microcentrifuge tubes and the remaining membrane fraction is resuspended in 500 µl 10 mM NaPi buffer. Both fractions are kept on ice while preparing for loading as described above.

Following electrophoresis, the resolved protein samples were transferred to PVDF membranes (MilliporeSigma, #IPVH00010) for immunodetection [78]. Immunoblotting of *Chlamydomonas* mitochondrial CI subunits was performed with antibodies raised against synthetic peptides as described previously [8]. Briefly, antibodies cross-reacting with mitochondrial proteins were used at the following conditions and dilutions: rabbit α-51 kDa (at 1:5,000), rabbit α-49 kDa (at 1:5,000), rabbit α-78 kDa (1:10,000), rabbit α-TYKY (1:10,000), and rabbit α-AOX1 (1:10,000). Antibodies cross-reacting with plastid resident proteins were used at the following conditions and dilutions: rabbit α-PsbO (1:9,000), rabbit α-PsbA (1:10,000), rabbit α-AtpA (1:8,000), rabbit α-AtpB (1:10,000), rabbit α-

RbcL (1:20,000), and rabbit α-cyt*f* (1:5,000). Mouse α-FLAG (Sigma-Aldrich, #F1804-50UG) was diluted 1:2,000 according to manufacturer's protocol. All primary antibody incubations were performed overnight at 4^{o}C with gentle agitation. The following antibodies were diluted in 0.1%TBS-T with 2% milk: α-FLAG, α-PsbO, α-PsbA, α-AtpA, α-AtpB, α-AOX1, α-cyt*f*, and α-RbcL. All others were diluted in 0.4%TBS-T with 5% milk. An HRP-conjugated goat anti-rabbit secondary antibody (Biorad, #170-6515) was used at 1:10,000 in the same buffer as for primary. For detection of mouse α-FLAG, a goat anti-mouse-HRP conjugated secondary was used (1:2,500 dilution, BIORAD, #1706516). All antibody dilutions were prepared in total volume of 25 ml.Five 5-minute washes after primary and secondary antibody incubations were performed. All washes consisted of 25 ml buffer (same % Tween-20 used for the primary antibody incubation). Bands of interest were visualized by incubation of the blots with a chemiluminescent substrate (ThermoFisher Scientific, #34095) and subsequent film exposure and development (Genemate, #F-9023-8x10).

4.3.8 Preparation of DNA, total RNA, and cDNA synthesis

Samples for DNA and RNA extractions were refreshed 2x on selective media before inoculation of 125 ml flasks containing 50 ml of TARG media to a starting OD_{750} of 0.2. Liquid cultures were incubated for 24 hrs then cells collected by centrifugation at 2,000 x *g* for 5 minutes. DNA and total RNA were extracted by phenol chloroform as described previously [78, 82] with several modifications. Following centrifugation, the

pellet was resuspended in 0.5 ml TEN buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 150 mM NaCl) and transferred to 1.5 ml microcentrifuge tubes. After removing the supernatant, the pellet was then resuspended in 150 µl ice cold ddH_2O and 300 µl of SDS-EB buffer (2% SDS, 400 mM NaCl, 40 mM EDTA, 100 mM Tris-HCl pH 8.0). Samples on ice were treated with Phenol/chloroform/isoamyl alcohol (PCI, 25:24:1 ratio, v/v) in a 1:1 ratio, shaken vigorously, and incubated at room temperature for 5 minutes with gentle agitation. After PCI treatment, the samples were clarified by centrifugation at 16,870 x *g* and the aqueous phase transferred to new 1.5 ml microcentrifuge tubes. A volume of 300 µl chloroform/isoamyl alcohol (CIA, 24:1 v/v) was then added to each sample, shaken vigorously, and incubated at room temperature with gentle agitation. The CIA treated samples were then clarified by centrifugation at 16,870 x *g*. This extraction method was repeated for a total of 2 rounds. Extraction of total cellular RNA was performed as for DNA but with phenol pH 5.0 (instead of pH 8.0 for DNA) and taking special precaution to prevent RNase contamination.

DNA was precipitated with ethanol as described previously [78] and treated with 50 µg RNase A (Amresco, #0675-250mg) to remove RNA. Extracted DNA was then quantified spectrophotometrically by nanodrop. Total RNA was precipitated by overnight incubation with 1/3 volume 8M LiCl at -20^{o}C and collected by centrifugation for 60 min at maximal speed and 4^{o}C. After the supernatant was removed, the RNA pellet was washed with 70% ethanol and re-collected by centrifugation at maximal speed for 20 minutes at 4^{o}C. The pellet was dried briefly at room temperature then resuspended in 50 µl RNase-free ddH_2O water and quantified by nanodrop. A total of 1-2.5 µg of RNA was resolved

electrophoretically and visualized by staining with ethidium bromide (Sigma-Aldrich, #E8751) to verify quality of extracts. Eight μg total RNA was treated with RQ1 RNase free DNase I (Promega, #M6101) to remove any residual DNA. DNA and total RNA extractions were performed in biological triplicates.

Eight μg of RNA was used in cDNA synthesis, performed by 400 units of M-MLV reverse transcriptase (Invitrogen, #28025-013) and 1 μg random hexamers (Promega, #C1181) according to manufacturer's protocols and as previously described [72, 132].

4.3.9 HPLC detection of total lipids

Starter cultures were inoculated as described above from recently refreshed strains on selective media and culture for 3-5 days with daily dilutions to maintain the cells in mid-exponential phase (180 rpm, constant light of 40-60 μmol photons m^{-2} s^{-1}, 25°C). Prior to inoculating sample cultures, cells were counted via hemocytometer, collected by centrifugation (5 min at 2,000 x *g*, 25°C), and resuspended to starting cell density of $2x10^6$ cells/ml in 35 ml TARG media (in 125 ml Erlenmeyer flasks). Cells were incubated in the above growth conditions for 24 hrs and OD_{750} measured prior to collection via centrifugation (5 min at 2,000 x *g*, 25°C). The pellet was resuspended in 1 ml media, transferred to 2.0 ml microcentrifuge tubes with locking cap (record weight of tubes for calculation of dry weight later), and collected again by centrifugation as described above. The supernatant was removed by pipetting or aspiration and flash frozen in liquid nitrogen

before storing at -80^{o}C. Samples were lyophilized and dry weight was measured before proceeding to total lipid extraction.

Extraction of total lipids was performed as done previously with several modifications [209]. First, the lyophilized cell pellet was resuspended in 1 ml lipid solvent (2:1:0.1, methanol: chloroform: formic acid) and vortexed for 5 minutes. A volume of 0.5 ml extraction buffer (1M potassium chloride, 0.2M phosphoric acid) was then added to each sample and vortexed for 20-30 minutes. Samples were centrifuged at 15,000 *x g* for 5 minutes (room temperature) and transfer bottom (organic) layer to new microcentrifuge tubes. The organic phase was transferred to new microfuge tubes and re-clarified by addition of 1x volume lipid dilution solvent (1:1, methanol: chloroform) and 0.5x volume lipid extraction buffer, followed by vortexing 2x 30 seconds, and then centrifuged for 5 minutes at 15,000 *x g*. The organic phase was transferred to new microfuge tubes as described above.

Prior to HPLC injection, samples were concentrated by speed-vac, resuspended in 150 µl lipid dilution solvent, and incubated at 42^{o}C for 5 minutes. To remove any precipitate that could clog the HPLC capillary tubing, samples were centrifuged at 15,000 *x g* for 1 minute and 120 µl sample transferred from top of sample to HPLC vials for injection via autosampler. Resolution of total lipids was performed using Alltech Altima C18 reverse phase column with 4.6 mm x 250 mm, 5 µm bead diameter dimensions (Alltech, #88056). A binary gradient from 15-100% solvent B (100% isopropanol) with reciprocal gradient of solvent A (90% acetonitrile, 10% 10mM ammonium acetate) was utilized with resolution time of 45 minutes (flow rate of 0.8 ml/minute) with 15 minute

wash between runs. Due to fluctuations in room temperature, the column was heated to 30^{o}C for standardization of runs elapsing several days. Lipid species were detected by absorbance at 215 nm. All solvents and reagents were HPLC grade or higher.

4.3.10 Detection of LCLA1 gene product

In silico manipulation of transforming DNA was performed using CloneManager and SnapGene programs (Figure 32). A unique *Sgr*DI restriction site was identified in pADC2 (*Bst*BI restricted 64B9 limited to *Cre07.g329861*) near the end of the *LCLA1* CDS (410 bp upstream of stop codon) and used for insertion of a synthesized DNA fragment encoding the 3xFLAG epitope tag (GeneWiz). The construct was designed such that it reconstituted the endogenous sequence adjacent to the *Sgr*DI site in pADC2 with flanking *Sgr*DI sites for subcloning into pADC2. One µg of purified pADC2 and the pUC57-Kan-3xFLAG (containing the sequence corresponding to the 3XFLAG) were digested with *Sgr*DI (ThermoFisher, #ER2031) for 1 hr at 37^{o}C then the reaction was stopped by heat inactivation according to manufacturer's instructions and resolved on 0.8% agarose gels. Linearized pADC2 and fragment corresponding to 3xFLAG tag were gel-extracted (Macherey Nagel, #740609.50) according to manufacturer's instructions and mixed (3:1 molar ration of insert:vector) and incubated overnight at 4^{o}C for relegation with T4 DNA ligase (Invitrogen, #15224017). Five µl of ligation reaction was then transformed into One Shot TOP10 cells (Invitrogen, #C404003) according to manufactures protocol and colony PCR used to screen for presence of the 102 bp insertion at the cloning site. Purified plasmid DNA was then prepared for positive clones (Machery Nagel NucleoSpin Plasmid,

#740588.250) following the manufacturer's protocol and presence of insert verified by DNA sequencing (See Table 4 for primer sequences).

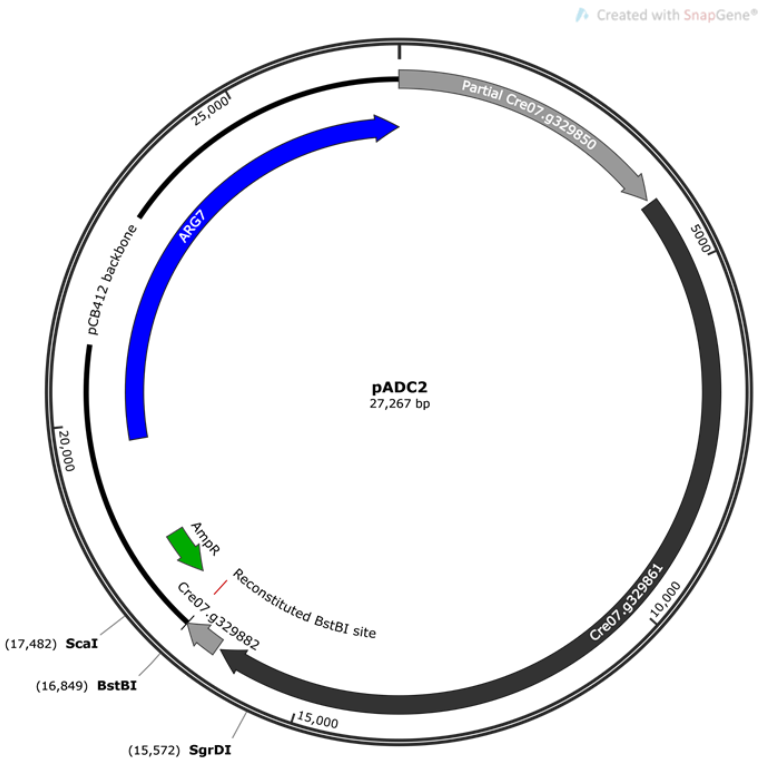


Figure 32 pADC2 recombinant vector map.

64B9 was restricted with *Bst*BI and re-ligated to yield pADC2, lacking the only other full gene (*Cre07.g329882*) present in the 64B9 cosmid. Vector map created with SnapGene Viewer.

The resulting recombinant vectors, pADC10 and pADC11, encoding *LCLA1-3xFLAG* and *LCLA1-6xFLAG* respectively (simplified as *LCLA1-FLAG*) were introduced into the *lcla1* (*lcla1#1*) mutant by glass bead transformation (described above). Transformants were selected based on arginine prototrophy, subcloned, and tested for rescue of the SID and high light sensitivity phenotypes. Complemented strains were then assayed for presence of the recombinant vector by genotyping PCRs and positive strains propagated ~2 weeks (refreshed 3x) in each of the indicated growth conditions before collection. Total protein extract was then collected from complemented *lcla1(LCLA1-3xFLAG)* transformants by freeze thaw method (described above) and resolved by SDS-PAGE on 4-10% gradient polyacrylamide gels (v/v) (16 x 16 cm, BIO-RAD, PROTEAN II xi cell system) for 5 hrs at 200V and transferred to PVDF membranes (Millipore Sigma, #IPVH00010) via Western transfer for 5 hrs at 100 V. Immunoblotting was performed as described above with the following exceptions: 1) membranes were blocked overnight in 100 ml 0.1%TBST containing 3% milk, 2) recombinant LCLA1-3xFLAG was detected by overnight incubation at 4^{o}C with a 1:2,000 dilution of α-FLAG antibody (Sigma Aldrich, #F1804), and subsequent 1 hour incubation at room temperature with goat anti-mouse-HRP conjugated secondary (1:2,500 dilution, BIORAD, #1706516). Both primary and secondary antibodies were prepared in 50 ml of 0.1%TBST and 3% milk. Washes (50 ml each), chemiluminescent detection, and film exposure were performed as described above.

4.3.11 Subcellular localization

In silico manipulation of transforming DNA was performed using SnapGene and CloneManager programs (Figure 33). Initial attempts utilizing a two-promoter system (one for transgene and second for selectable marker) were unsuccessful. Indeed, difficulties regarding transgene expression in *Chlamydomonas* are well established, including random mutation, positional control of gene expression, fragmentation of transgene, and rapid silencing of transgene expression [90]. Therefore, an alternative parent vector (pMO449) utilizing a one-promoter system in which expression of the selectable marker was transcriptionally linked to that of the transgene was chosen [206]. The sequence encoding the N-terminal 72 amino acids of LCLA1 was amplified using gene specific primers (LCLA1-S-N72-F and LCLA1-S-N72-R) from random hexamer primed cDNA (See Table 4 for primer sequences), inserted via a unique *Hpa*I restriction site into the pMO449 fluorescent reporter expression vector [206] (obtained from the *Chlamydomonas* Resource Center (https://www.chlamycollection.org/), and transformed into chemically competent DH5α bacteria. Presence and sense-orientation of the inserted PCR product was diagnosed by colony PCR (See Table 4 for primer sequences). Plasmid DNA was purified from positive clones according to manufactures instructions (Machery Nagel NucleoSpin Plasmid, #740588.250) and confirmed by DNA sequencing.

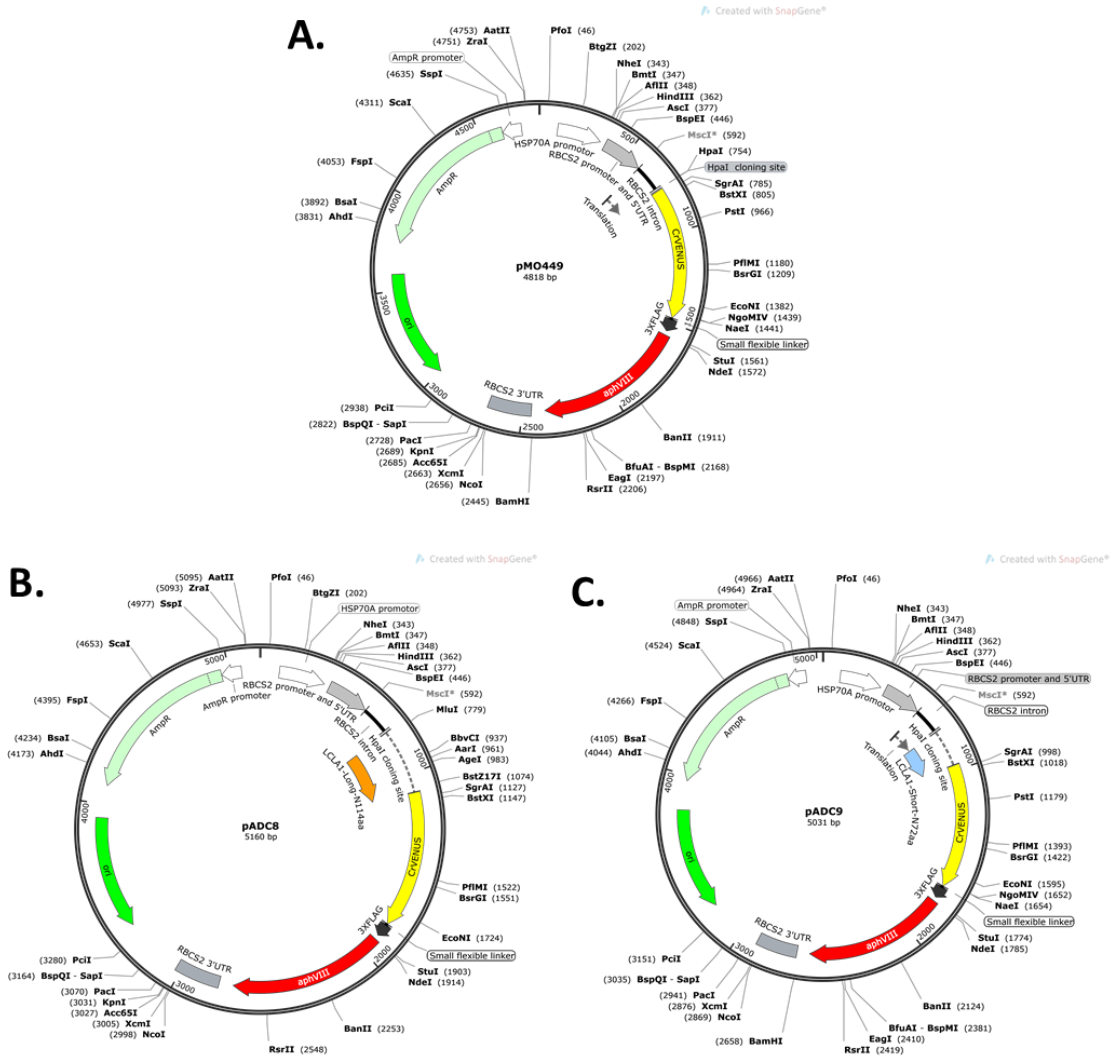


Figure 33 Vectors for expression of LCLA1 chimeric fluorescent reporter.

A) Map of the pMO449 expression vector obtained from the *Chlamydomonas* Resource Center. Recombinant vectors were generated by blunt-end cloning via *Hpa*I site (grey highlighting). B) $LCLA1^{L}$-N114-CrVENUS recombinant vector designated as pADC8. C) $LCLA1^{S}$-N72-CrVENUS recombinant vector designated as pADC9. All unique restriction sites shown; the *Sca*I site present in the *bla* gene imparting ampicillin resistance was used to linearize the vector for improved transformation efficiency. Vector maps (not to scale) made using SnapGene Viewer.

The resulting vector, pADC9 (or pMO449 empty vector) was linearized by restriction digest with *Sca*I or *Eco*RV and introduced into WT *Chlamydomonas* (CC-4533 or CC-5155) by glass bead transformation as outlined above. Since expression of the fluorescent reporter is transcriptionally linked to that of the antibiotic resistance marker (*APHVIII*) in this vector, prescreening of transformants by relative fluorescence (via microplate reader) was unnecessary. Pm^R colonies were secondarily screened by PCR and positive transformants subcloned. Three independent transformants were selected based on these pre-screening criteria and subcellular localization investigated by confocal fluorescent microscopy (described below) compared to known subcellular markers [201] and organelle specific dyes.

4.3.12 Organelle staining

Samples of interest were refreshed 1-2x on selective media before inoculation of 200 ml TAP or TARG media in 500 ml Erlenmeyer flasks and grown as described above for 3-5 days with daily dilutions prior to staining and confocal fluorescence microscopy (described below). For staining, cultures were collected by centrifugation at 1,000 x *g* for 5 minutes (room temperature) and resuspended to a final cell density of $1x10^8$ cells/ml with 1 ml of each sample transferred to 2.0 ml microcentrifuge tubes. Mitochondria were stained with Mitotracker Red CMXRos (Lifetech, #M7512) dissolved in DMSO (1 μM final concentration) for 30 minutes in the dark with gentle agitation. Following stain incubation, cells were collected by centrifugation at 100 x *g* for 5 minutes at room temperature, washed with 1 ml fresh TARG media, recollected by centrifugation, and resuspended in 200 μl. A

volume of 2.5 µl of each sample were then added to microscope slides (Fisher, #12-544-1) containing 2.5 µl of molten TAP+1.5% low-melting agarose cooled to 37^{o}C (pipetting to evenly mix) and coverslip applied. To prevent desiccation of samples, the coverslips were then sealed with nail polish. Staining of lipid bodies was performed analogous to Mitotracker dye but with Nile Red (Sigma-Aldrich, #72485) with final concentration of 10 µg/ml (31.4 µM in 100% ethanol) per 1x10^{8} cells/ml (1 ml total volume in 2.0 ml microfuge tubes). Staining of lipid bodies with Nile Red was performed in $\geq$3 biological replicates.

4.3.13 Expression of fluorescent markers in the *lcla1* mutant

The indicated fluorescent reporters [201] expressed in the transgenic strains were introduced into the *lcla1* mutant by genetic crosses (described above, see Table 3 for strains). Recombinant spores were isolated by selection on TAP+Hyg+Pm and subcloned to single colonies. Six independent haploid progeny were investigated for each fluorescent marker expressed in the *lcla1* mutant. Indicated strains were refreshed 1-2x from stocks on selective media then a loop-full of 2-5 day old grown cultures on plates was used to inoculate 35 ml TAP media in 125 ml Erlenmeyer flasks. Strains were cultured as described above for 3-5 days with daily dilutions prior to confocal fluorescencemicroscopy (described below).

4.3.14 Confocal fluorescence microscopy

Sample preparation for strains expressing fluorescent reporters (i.e. not stained with organelle specific dye) was performed differently than described above. Indicated strains were refreshed 1-2x on selective media then used to inoculate 35 ml TAP media in 125 ml Erlenmeyer flasks and maintained by daily dilutions in an OD_{750} range of 0.2-0.6 (40-60 μmol photons m^{-2} s^{-1} constant illumination unless otherwise indicated, 25ºC, 180 rpm). One ml culture was then collected from each sample by centrifugation at 2,000 x *g* for 1-2 mins at 25ºC and the supernatant removed by pipetting or aspiration. Cell pellets were resuspended with molten TAP media containing 1.5% low-melting agarose cooled to 37ºC and spotted onto microscope slides (Fisher, #12-544-1), and the added coverslip sealed by nail polish to prevent desiccation. Live samples and controls were observed by spinning disc confocal microscopy (Nikon Eclipse Ti inverted microscope equipped with EMCCD C9100-23B camera) at 600x total magnification. CrVENUS was detected by excitation at 514 nm and chlorophyll autofluorescence at 640 nm (both lasers set to 20% laser power and 10-20% sensitivity). MitoTracker Red CMXRos was detected by excitation at 561 nm whereas Nile Red was detected at 514 nm. Bright-field images were collected by differential interference contrast (DIC). Images were processed using the Volocity program (Perkin Elmer, version 6.3.1).

Nile Red staining bodies were counted per cell for each sample. Size of lipid bodies was inferred by total fluorescence in marked area measured by Volocity program. Variation in total number of lipid bodies and lipid body size was displayed by box and whisker plot (> 150-250 total cells from $\geq$ 3 replicates).

Mitochondrial morphology was scored as reticulate, lobed, punctate/fragmented, or swollen in 2 independent, haploid progeny compared to the parental FP expressing strain. Percent of each category graphed based on average of ≥ 4 independent replicates of 50+ cells with error bars representing standard deviation.

4.3.15 Bioinformatic prediction of subcellular localization and secondary structure

Subcellular localization of LCLA1-Long and LCLA1-Short was predicted by comparison to proteins of known subcellular localization using the following prediction programs with the default settings: TargetP 1.1 [145], PredAlgo [146], Mitofates [147], TargetP 2.0 [144], DeepLocP [210], SherLoc2 [211], and WoLF PSORT [212] (summarized in Table 6). Presence of transmembrane domains ™ were predicted using the following programs: Das [213], TMpred, and TMSEG [214] and mapped onto each LCLA1 isoform. Tm prediction was then compared to secondary structure predictions using GOR [215]. The EMBOSS "stretcher" algorithm (https://www.ebi.ac.uk/Tools/psa/emboss_stretcher/) was used for global pairwise sequence alignment of large protein sequences and compared to the "water" algorithm (https://www.ebi.ac.uk/Tools/psa/emboss_water/) for local alignments [216]. The SMART program (https://smart.embl.de/) was used to identify low complexity regions in LCLA1-Long and LCLA1-Short [217, 218]. MEME was used to identify short shared motifs in related protein sequences and FIMO used to identify other proteins with each motif (https://meme-suite.org/meme/tools/meme) [168, 191]. Models for chimeric proteins, *etc*. were created with BioRender.com.

4.4 Results

4.4.1 Cre07.g329861 is required for growth in respiratory and mixotrophic conditions

Prior genetic analysis of the *amc12(#6E9)* strain had identified one insertional mutation in the *Cre07.*g329861 gene and a second, independently segregating nuclear mutation (*sid2*) unlinked to the insertional cassette (unpublished, data not shown) [72]. The presence of a second mutation not due to insertion of the antibiotic resistance cassette was not entirely unexpected as the SID phenotype in 8 of the 13 *amc* mutants is not linked to the insertional cassette [8, 72, 132]. For example, an insertion of *Chlamydomonas* chromosome 17 was found to disrupt the *AMC1* gene in *amc1-1*. However, that this second mutation also contributed to the SID phenotype was not anticipated. Isolation of the single mutation, in an otherwise wild-type background, preceded any investigation of *Cre07.g329861* (hereafter *LCLA1*). To this end, Hyg^R haploid progeny resulting from sexual crosses of the *amc12* original mutant to wild-type 3 additional times for a total of 4 backcrosses. The wild-type strain chosen for these crosses was CC-4533, used in production of the *Chlamydomonas* Library Project (CLiP) [11, 202], and its isogenic mating partner CC-5155. The reasons for choice this were three-fold: 1) each wild-type strain had very robust growth in the dark facilitating easy selection of meiotically recombined SID progeny, 2) introgression in this background would enable direct comparison of the *lcla1* mutant phenotype to other strains originating from this library, and 3) initial attempts to backcross the *amc12* original mutant to its compatible wild-type

mating partner were unsuccessful; however, crosses with this background were very efficient as has been reported before [11, 202].

To address whether disruption of *LCLA1* was causative for the SID phenotype observed in the original mutant, tenfold dilution series was performed after the 4th (final) backcross of 8 independent haploid progeny. In addition, one of these haploid progeny was selected for an additional backcross for tetrad analysis (Figure 34). As expected, all Hyg^{R} haploid progeny were SID confirming that the SID phenotype was genetically linked to *LCLA1*. To address whether loss of LCLA1 alone was directly responsible for the SID phenotype and not a mutation present in a nearby gene which would co-segregate with the antibiotic resistance marker disrupting *LCLA1*, we sought to complement the *lcla1* mutant by introducing the wild-type copy of the gene. To this end, we took advantage of an indexed cosmid-based genomic library containing the entirety of the *Chlamydomonas* genome cloned in 20-40 kb segments, with an average of 38 kb per segment [76, 205]. The 64B9 cosmid was identified by PCR-based screening of this library with primers specific to the *LCLA1* gene (Figure 30) and introduced into the *lcla1* mutant via glass bead transformation. As expected, transformation with the WT copy of *LCLA1* contained in 64B9 complemented the SID phenotype. Since a second full gene was contained within the 64B9 cosmid, this vector was further restricted with *Bar*I disrupting the downstream gene, *Cre07.g329882*, or with *Sgr*DI disrupting *LCLA1* and introduced into *lcla1* (Figure 31). As expected, 64B9 in which *Cre07.g329882* was restricted but not *LCLA1* was able to complement the SID phenotype in the *lcla1* mutant (Figure 31C). Similarly, transformation of *lcla1* with the empty vector (pCB412) did not rescue the SID phenotype (Figure 30A).

The *Bar*I restriction site in *Cre07.g329882* is ~5.4 kb downstream of the end of *LCLA1*. To further limit the transforming DNA required for complementation, 64B9 was shortened by restriction digest. The modified cosmid, pADC2 was introduced into the *lcla1* mutant and transformants tested for rescue of the SID phenotype by tenfold dilution series. As expected, this version of the cosmid containing only the wild-type copy of *LCLA1* was able to complement the respiratory growth deficiency on solid media (Figure 35).

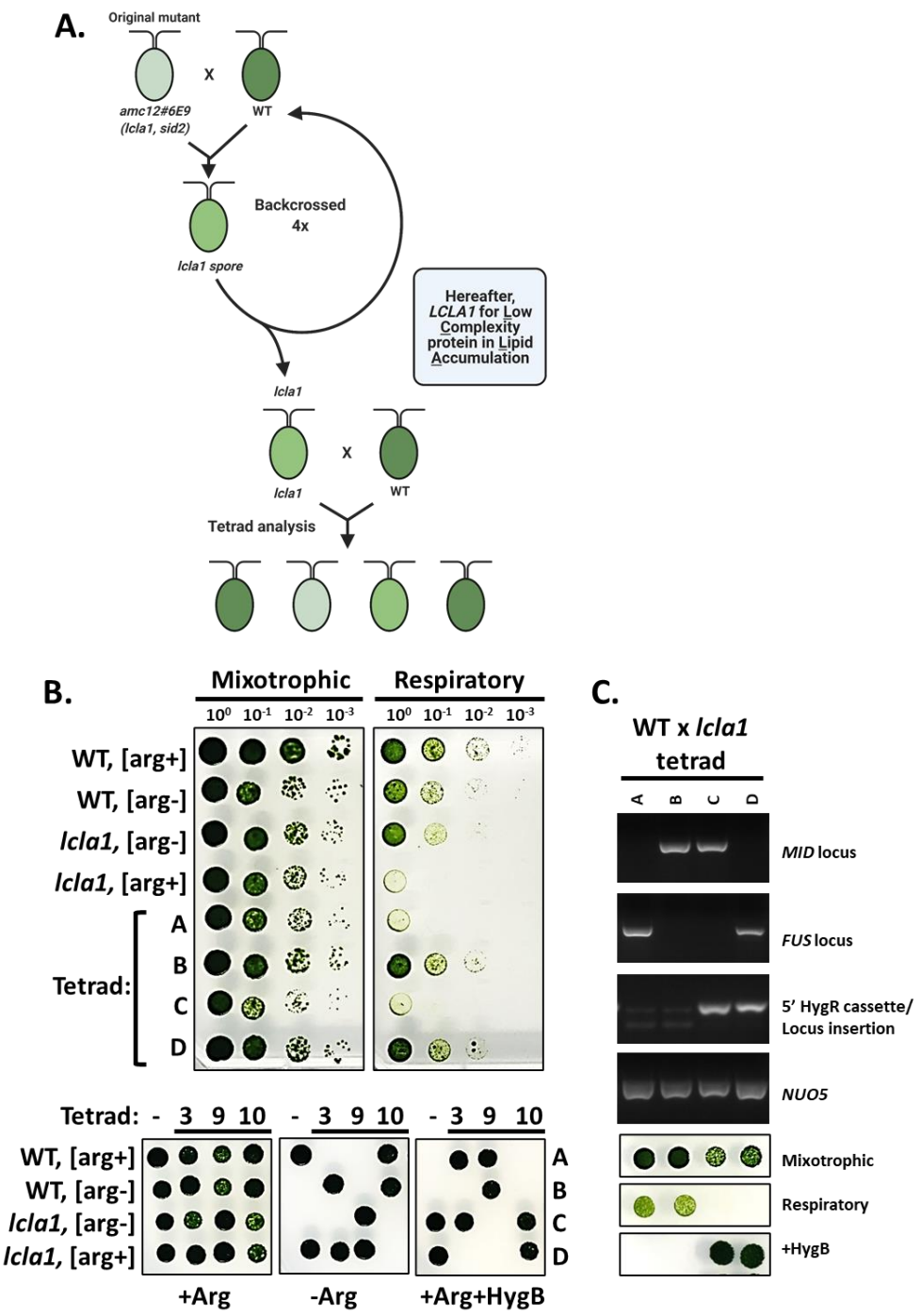


Figure 34 Isolation of single lcla1 mutant with SID phenotype.

A) Schematic for backcrossing strategy for isolation of *lcla1* single mutant and tetrad analysis. WT is CC-4533 and CC-5155 were used in backcrossing (isogenic opposite

mating partners) depending on mating type of Hyg^R spore selected. B) Tenfold dilution series of representative tetrad and control strains (top panels) at 7 days of indicated growth condition (of 9 tetrads tested). B) Lower panels, first column control strains indicated for markers, remaining 3 columns are representative tetrads resulting from the cross of *icl1*, [arg-] x WT, [arg+] with segregation of [arg-] and Hyg^R selectable markers. C) PCR-based genotyping of a representative tetrad showing 2:2 segregation of genetic markers (top 3 panels) compared to amplification of region of NUO5 gene (4^{th} panel from top). C) Single spots at 10^{-1} dilution show co-segregation of genotypic and phenotypic markers (lower 3 panels of B and C).

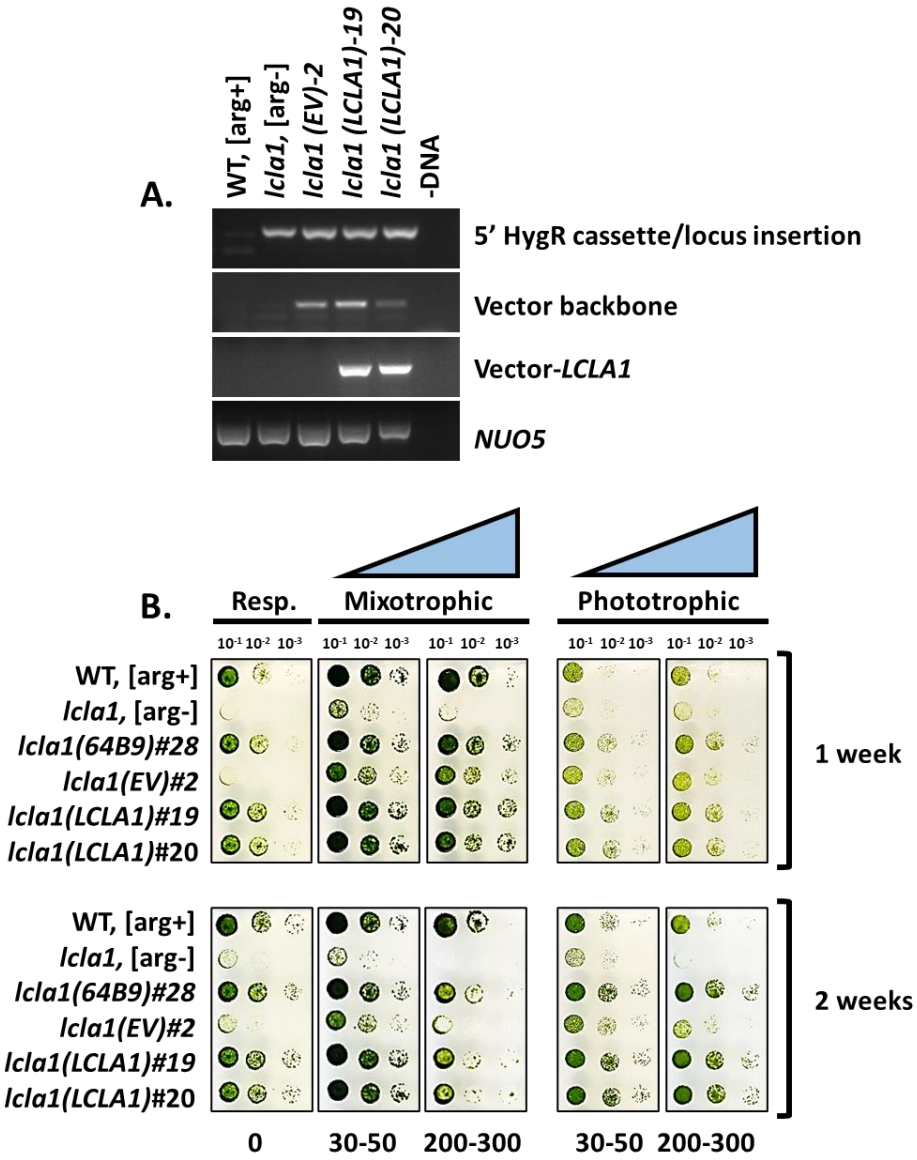


Figure 35 The *lcla1* insertional mutant is complemented by the *LCLA1* gene.

A) PCR genotyping of indicated strains tested for growth in (B) with the following primer pairs: 5' iHyg3 insertion (6E9-RT2-F1/APH7-R5), vector backbone (PH1/PH2), vector-*LCLA1* (PH3/64B9-480R), and *NUO5* (NUO5-E2L/NUO5-E3R). See Table 4 for primer sequences. B) Tenfold dilution series of the indicated strains. WT, [arg+] is CC-4533, *lcla1*, [arg-] is l*cla1#1*. Images for top panel taken at 1 week whereas bottom panel taken at 2 weeks. Light intensity is increasing from left to right with range (µmol photons m^{-2} s^{-1}) described at bottom. Minimal media is same recipe except lacking acetate (recipe details

in methods). Transformed strains (with parentheses) are all [arg+] due to *ARG7* selectable marker of cosmid (see Table 5).

The *LCLA1* gene is currently annotated in the Phytozome browser as *Cre07.g329861.t1.2*. *Chlamydomonas* gene models are predicted based on the GreenGenie2 algorithm and confirmed based on an abundance of transcriptomic data and manual validation [2-4]. However, initial attempts to amplify the 5' end of the transcript containing the start codon via 5'-RACE or RT-PCR were unsuccessful whereas the 3' end of the transcript was readily detectable (data not shown). These results lead us to hypothesize that the gene model may not be fully accurate. Additional sequence information from the Merchant group (Dr. Sean Gallaher, personal communication) revealed the *LCLA1* gene was alternatively spliced at the 5' end of the transcript with neither of the two spliced transcripts (*Cre07.g329861.t2.*2 and *Cre07.g329861.t2.*) matching the current version of the predicted transcript present on the Phytozome genome portal. *Cre07.g329861.t2.2*, the longer of the two splice variants has a transcript of 10,719 nts with coding sequence of 10,041 nts (predicted to encode 3,347 amino acid protein) whereas the transcript of the shorter splice variant, *Cre07.g329861.t2.1*, is 10,561 nts with coding sequence of 9,675 nts (predicted to encode 3,224 amino acid protein). Importantly, based on transcriptomic data, these two splice variants vary only at the 5' end. Therefore, the encoded proteins would be expected to be identical in sequence with exception of the N-terminal extension encoded by the longer transcript. RT-PCR of the 5' end of the transcript with primers designed based on the revised gene model were performed and the resulting amplicons cloned. DNA sequencing of the cloned amplicons verified the 5' end of both transcripts *Cre07.g329861.t2.1* and *Cre07.g329861.t2.2* predicted by Dr. Gallaher's analysis (Figure 29B-C).

During prior growth analyses by tenfold dilution series, it was observed that the *lcla1* mutant grew more slowly than the corresponding wild-type and complemented strains in mixotrophic conditions (light plus acetate, Figure 35B). To quantitatively assess the difference in growth rate of the *lcla1* mutant compared to wild-type, empty vector-transformed, and complemented strains, growth curve analyses were performed. The doubling time of *lcla1* (~13 hours) was approximately twice that of wild-type (~7 hours) in mixotrophic conditions (Figure 36A-B,E). Transformation of *lcla1* with the empty vector did not improve doubling time in mixotrophic conditions (~22 hours) whereas introduction of the wild-type copy of *LCLA1* rescued the growth deficiency to wild-type levels (~7 hours) (Figure 36A-B,E). Likewise, growth of the *lcla1* mutant in respiratory conditions was slower (~17 hours) compared to wild-type (11 hours) (Figure 36C-D,E). As expected, transformation of *lcla1* with the empty vector did not improve growth (29 hours) whereas introduction of the wild-type copy of *LCLA1* complemented the respiratory growth deficiency (~11 hours) in liquid media (Figure 36C-D,E). Calculation of doubling time was similar whether determined from change in OD_{750} or calculated cell density. Maximum OD_{750} was equivalent for all strains in mixotrophic or respiratory conditions (Figure 36F). Unexpectedly, maximum cell density was not the same across conditions and samples as was the case for maximum OD_{750} (Figure 36G). Taken together, these results indicated that *LCLA1* alone was responsible for the growth defect in respiratory and mixotrophic conditions.

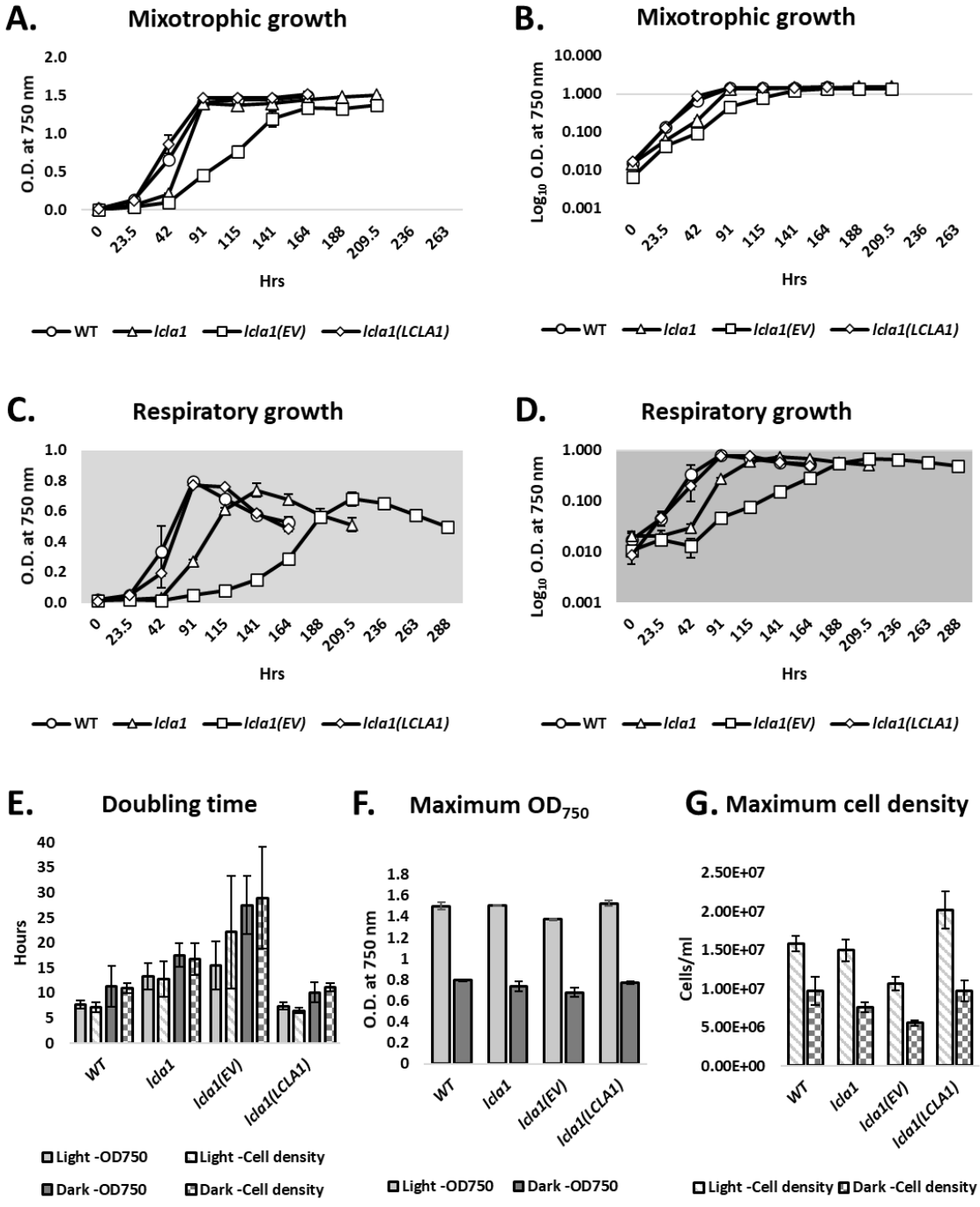


Figure 36 Doubling time of *lcla1* is longer in mixotrophic and respiratory conditions.

Change in growth of *Chlamydomonas* culture inferred from optical density (O.D.) at 750 nm in mixotrophic (A and B) and respiratory (C and D) plotted as a function of time.

Growth curves were log-transformed (B and D) for calculation of doubling time plotted in E. Cell density was also observed as a function of (growth curves not shown) and doubling time calculated, plotted in E. F) Maximum O.D. and cell density were recorded as a measure of total growth. WT, [arg+] is CC-4533, *lcla1* is *lcla1* spore #1, *lcla1(EV)* is #2, and *lcla1(LCLA1)* is #20. TARG media was used for all strains and in both light conditions.

4.4.2 Acetate is not limiting for growth of the *lcla1* mutant

Since it was observed that the *lcla1* mutant grew slower in liquid media and on solid media (appearing as petite colonies) in mixotrophic conditions, we reasoned that an ingredient required for growth could be limiting. To date, acetate is the only carbon source known to support extended growth of wild-type *Chlamydomonas* cultures in respiratory conditions [1, 219]. In mixotrophic conditions, acetate is also assimilated in addition to carbon from atmospheric CO_2 which occurs via photosynthesis [1]. Since the *lcla1* mutant was most noticeably affected for growth on solid media in conditions requiring acetate assimilation (i.e. dark), we asked whether acetate was limiting for growth. To address this question, the *lcla1* mutant was spotted via tenfold dilution series on plates containing increasing concentrations of acetate compared to control strains. To control for differences in osmolarity which would occur when increasing the concentration of acetate, tenfold dilution series was also performed where sorbitol, a non-utilizable carbon source was supplemented to the same molar concentrations. Increasing acetate concentrations had no effect on growth in either mixotrophic or respiratory conditions compared to media supplemented with sorbitol (data not shown). We concluded acetate is not limiting the respiratory and mixotrophic growth in the *lcla1* mutant.

4.4.3 Growth of *lcla1* in mixotrophic conditions is modulated by the *ARG7* nutritional marker

Because the cosmid used for the construction of the library was designed with the *ARG7* marker [76, 205], the *lcla1, arg7-8* haploid progeny had been retained for selection of transformants carrying the cosmid based on arginine prototrophy. Surprisingly, empty-cosmid transformed *lcla1* (*lcla1, arg7-8; ARG7*) grew better in mixotrophic conditions on solid media than the untransformed recipient strain (Figure 35B). Likewise, *lcla1,* [arg+] haploid progeny grew better than the corresponding [arg-] haploid progeny obtained from the same cross whereas no growth difference was observed between wild-type, [arg+] or [arg-] haploid progeny (Figure 37 and 38B and not shown). Similarly, no difference was observed between *amc9,* [arg+] and [arg-] haploid progeny (Figure 37 and not shown).

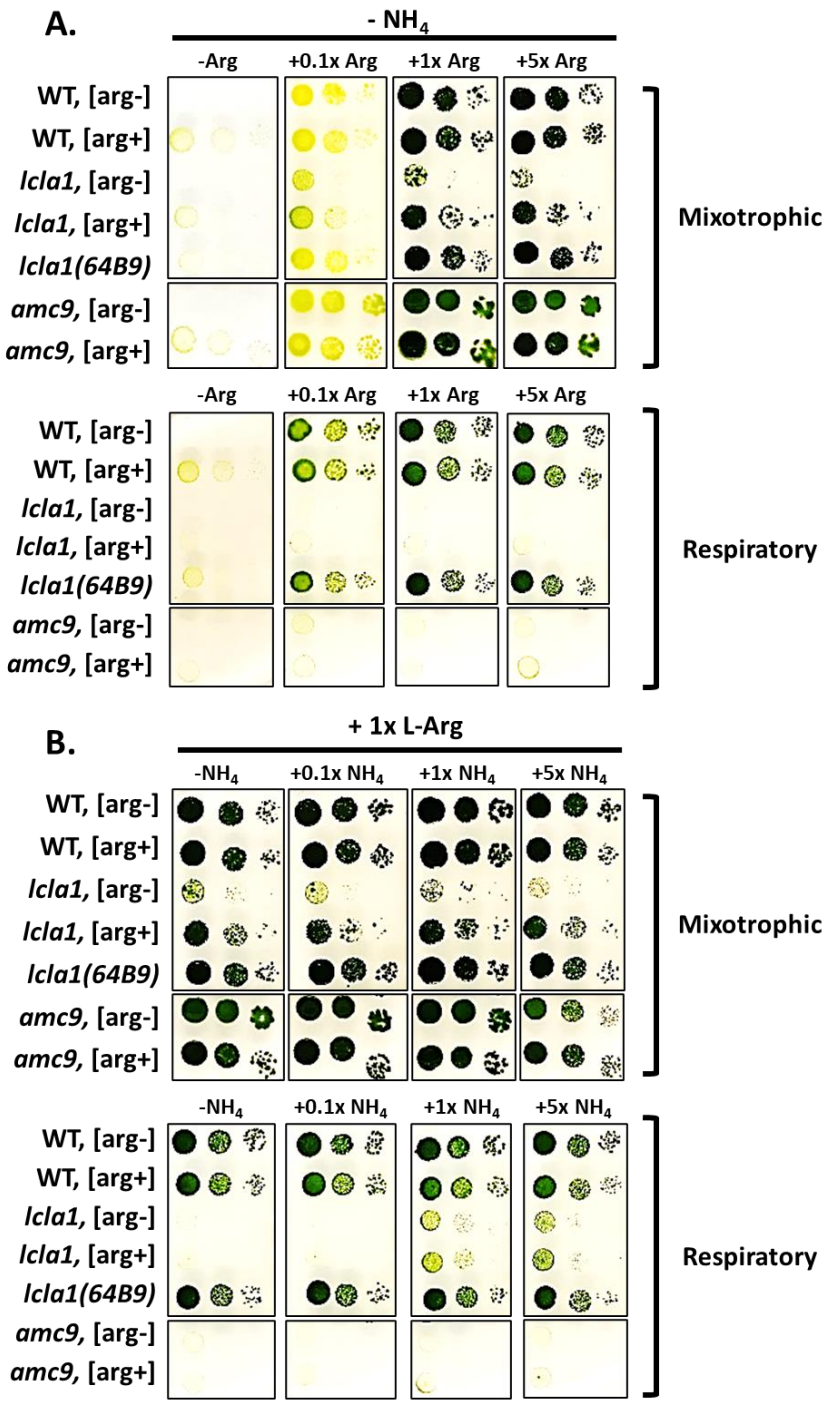


Figure 37 *lcla1* cannot utilize L-arginine as a nitrogen source in respiratory conditions.

Tenfold dilution series of indicated strains in media lacking ammonium (the preferred nitrogen source) with increasing concentrations of L-arginine (A) compared to media with

increasing concentrations of ammonium (B). L-arginine is required for arginine auxotrophic strains WT, [arg-], *lcla1*, [arg-] and *amc9*, [arg-] but each can utilize L-arginine as a nitrogen source in mixotrophic conditions (A, top panel) similar to ammonium (B, top panel). In the dark, *lcla1*, [arg+] and *lcla1*, [arg-] cannot grow with L-arginine as the sole nitrogen source (A, bottom panel), even when supplemented at 5x the normal concentration but can grow in the dark when ammonium is provided (B, bottom panel). Arginine prototrophic strains used: WT (CC-4533), *lcla1#5*, *amc9PH39-3*. Arginine auxotrophic strains used: WT spore #1, *lcla1 spore #1*, and *amc9#PH41-27*. All pictures taken at 10 days in the light or dark. Cultures in the light (mixotrophic) exposed to light intensity of ~ 30 µmol photons $m^{-2}s^{-1}$ or 0 µmol photons $m^{-2}s^{-1}$ for respiratory growth.

Chlamydomonas ARG7 (*Cre01.g021251*) encodes arginino-succinate lyase, catalyzing the final step of L-arginine biosynthesis producing L-arginine and fumarate from L-arginino-succinate [220]. It has previously been established that L-arginine can support growth *Chlamydomonas* as the sole nitrogen source [1, 221]. Indeed, it is one of a few amino acids uptaken and the only one done so with high affinity (with other nitrogen rich amino acids, histidine and glutamine less so). Other amino acids (including alanine, asparagine, cysteine, isoleucine, leucine, lysine, methionine, phenylalanine, serine, and valine) can serve as the sole nitrogen source but do not enter the cell; instead an extracellular deaminase such as *Chlamydomonas* LAO1 cleaves off the amine group, converting it to ammonia which is then imported [222-224]. In contrast, utilization of L-arginine is independent of LAO1 activity as growth of the *lao1* mutant is equivalent to wild-type with L-arginine as the sole nitrogen source [224]. To investigate whether L-arginine or nitrogen were limiting for growth of the *lcla1* mutant, tenfold dilution series was performed on N-dropout media supplemented with increasing concentrations of arginine or NH_4, the preferred nitrogen source. In mixotrophic conditions, wild-type, *lcla1*, and control strains grew normally with either nitrogen source (Figure 37). However, in respiratory conditions, L-arginine at 5-fold normal concentration was unable to support growth of *lcla1*, [arg+] and [arg-] spores in the absence of NH_4 whereas the respective wild-type and *amc9* strains grew normally (Figure 37A, lower panel). In contrast, supplementation with NH_4 and L-arginine enabled growth of *lcla1* in respiratory conditions albeit still less than wild-type, [arg+] and [arg-] spores (Figure 37B).

To further probe the connection between L-arginine biosynthesis and growth of the *lcla1* mutant, a pharmacological approach was taken. In the yeast model system *Saccharomyces cerevisiae*, arginine deficient cells express the Can1p arginine transporter localized to the plasma membrane to facilitate arginine uptake from the media. However, when arginine levels are sufficient in the cell interior, the Can1p transporter is internalized via endocytosis to the yeast vacuole mediated by Rsp5 and ARTs (Arrestin-Related Trafficking) adaptors [225] (Figure 38A). Can1p is so named because yeast mutants defective in L-arginine biosynthesis were sensitive to canavanine, an amino acid structurally similar to L-arginine which becomes cytotoxic when incorporated in place of arginine during translation (due to production of misfolded proteins) [226, 227]. To date, we are aware of only a few studies detailing the effects of canavanine toxicity in *Chlamydomonas* [228, 229] but neither thoroughly documented its effect on growth.

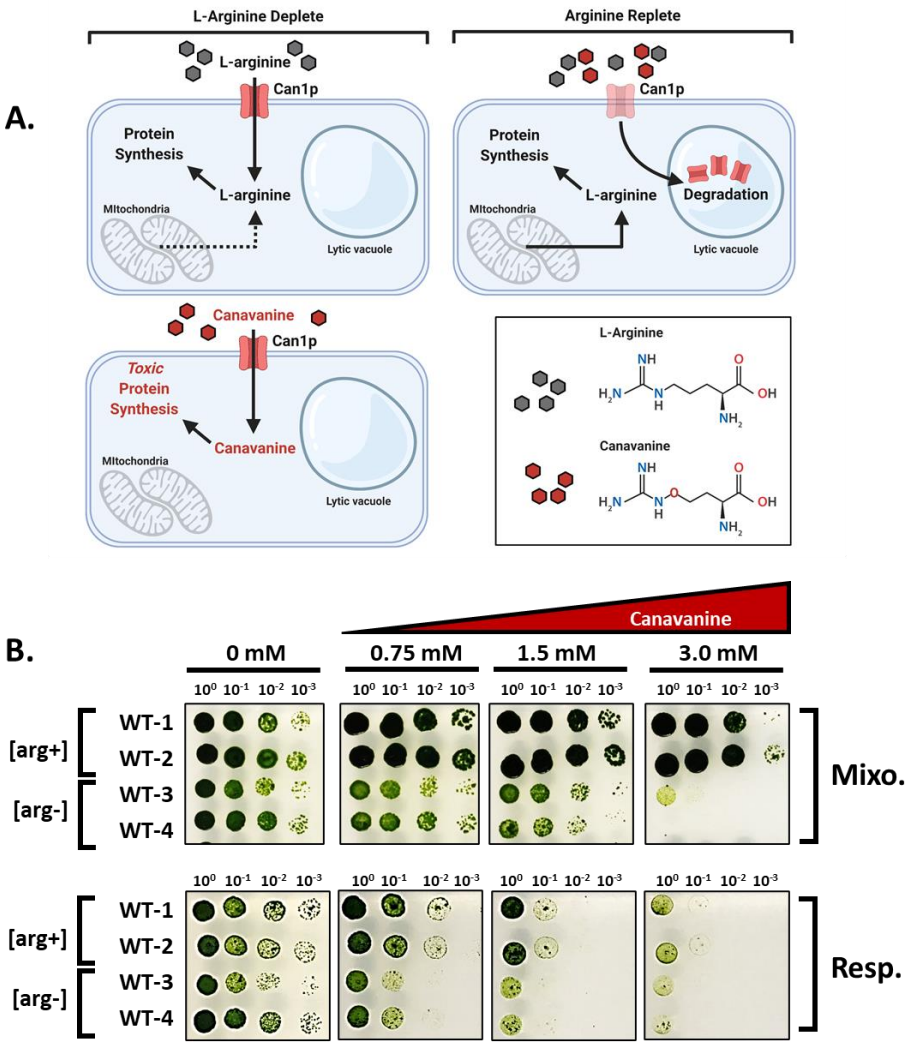


Figure 38 *Chlamydomonas* arginine auxotrophic strains are sensitive to canavanine.

A) model of canavanine toxicity. B) Tenfold dilution series of spores of the same background incubated in the light (30 µmol photons $m^{-2}s^{-1}$) or in the dark for 10 days in the presence of increasing concentration of canavanine. L-arginine was supplemented in the media at 0.04mg/ml L-Arg (one tenth of the normal concentration), all other ingredients for TARG as outlined in methods. Arginine prototrophic strains: WT-1 (CC-4533) and WT-2 (CC-5155). Arginine auxotrophic strains: WT-3 (WT#PH25-1) and WT-4 (WT#PH25-2). Model in (A) created using BioRender.com.

Growth of wild-type *Chlamydomonas* was compared for [arg+] versus [arg-] strains on solid media supplemented with increasing concentrations of canavanine relative to L-arginine in mixotrophic and respiratory growth conditions. As expected, wild-type [arg-] strains were more sensitive to treatment with canavanine compared to [arg+] strains in mixotrophic conditions (Figure 38). Interestingly, wild-type [arg-] and [arg+] strains were more sensitive in respiratory growth conditions than in the light suggesting an increased need for this particular amino acid (or its catabolite derivatives) for sustained culture in the dark (Figure 38). In comparison to wild-type and *amc9* haploid progeny of the same background, *lcla1* haploid progeny were hyper-sensitive to canavanine which was exacerbated genetically by combination with *arg7-8* allele and in the dark (Figure 39-40). This was also observed for *icl*, [arg-] spores in the light (data not shown). In the light, *lcla1* [arg+] haploid progeny were only sensitive at higher concentrations of canavanine (data not shown). In summary, these results suggest that growth of the *lcla1* mutant was limited by L-arginine even though the concentration in normal TARG media is >10-fold excess required for growth.

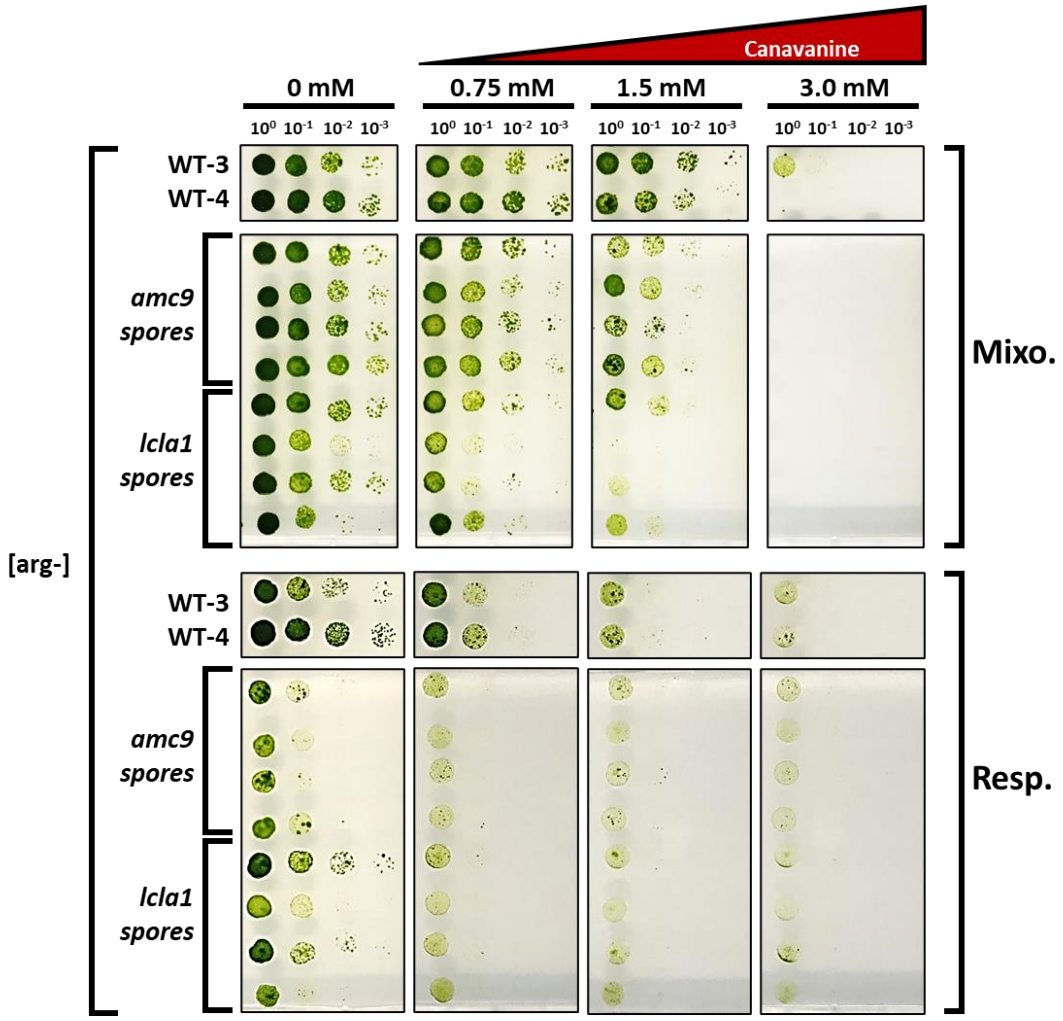


Figure 39 *lcla1*, [arg-] spores are hypersensitive to canavanine.

Tenfold dilution series of spores of the same background in the presence of increasing concentration of canavanine incubated in the light (top panel, 30 µmol photons $m^{-2}s^{-1}$) or in the dark (bottom panel) for 10 days. L-arginine was supplemented in the media at 0.04mg/ml L-Arg (one tenth of the normal concentration), all other ingredients for TARG as outlined in methods. Wild-type strains used: WT-3 (WT#PH25-1) and WT-4 (WT#PH25-2). *amc9* spores used: *#PH42-16*, *#PH39-26*, *#PH40-11*, and *#PH41-27*. *lcla1* spores used: *lcla1*#1 through #4. All strains shown are auxotrophic for arginine (contain *arg7-8* allele).

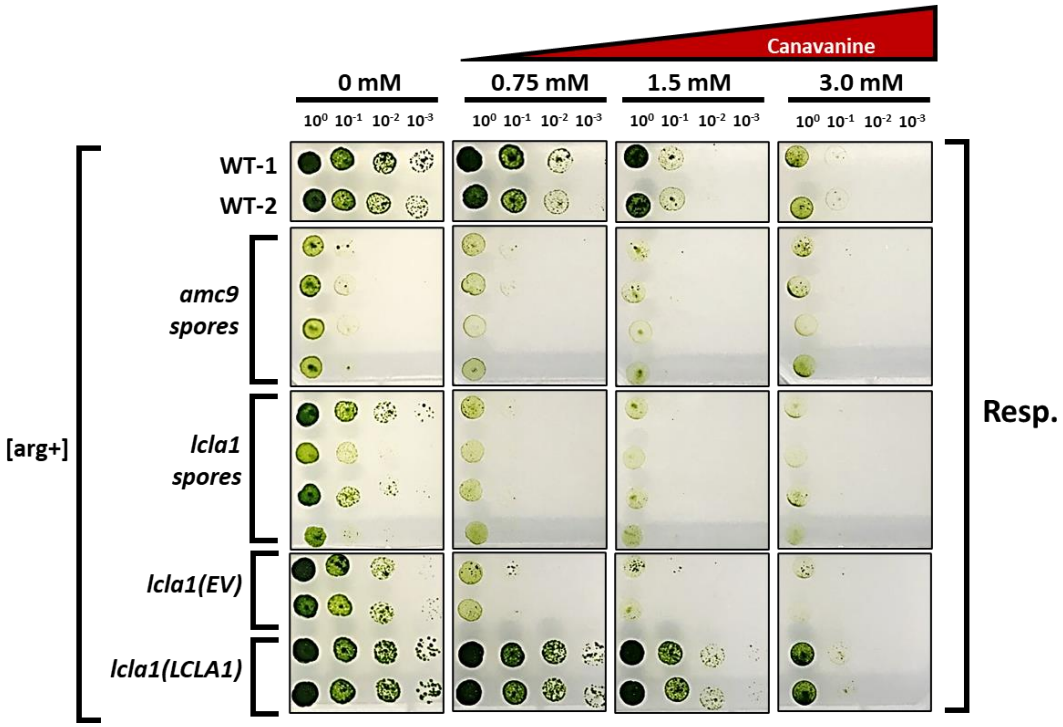


Figure 40 *lcla1*, [arg+] spores sensitive to canavanine in respiratory conditions.

Tenfold dilution series of backcrossed spores of the same background incubated in the dark for 10 days in the presence of increasing concentration of canavanine. Arginine was supplemented in the media at 0.04mg/ml L-Arg (one tenth of the normal concentration), all other ingredients for TARG as outlined in methods. Strains used: WT-1 (CC-4533), WT-2 (CC-5155). *amc9* spores used: *#PH40-31, #PH42-48, #PH41-18,* and *#PH39-3*. *lcla1* spores used: *lcla1*#5 through #8. *lcla1(EV)* are transformants #1 and #2. *lcla1(LCLA1)* complemented strains are transformants #19 and #20. All strains shown here are arginine prototrophic.

4.4.4 The *lcla1* mutant is sensitive to high light

Excess arginine was unable to rescue the growth deficiency of *lcla1* in mixotrophic or respiratory growth conditions whereas introduction of the *ARG7* gene via crosses or through transformation of *lcla1,* [arg-] haploid progeny only partially alleviated the mixotrophic growth phenotype and had no effect on growth in the dark (Figure 35B and 41). An alternative possibility for the observed slower growth in mixotrophic conditions was a defect in the photosynthetic electron transport chain. Photosynthetic mutants are often unable to grow on minimal media lacking acetate as a carbon source and can also be sensitive to high light [9, 10, 230]. In high light conditions, electrons cannot efficiently pass through the photosynthetic electron transport chain, escape, and react with molecular oxygen to produce singlet oxygen, a damaging reactive oxygen species [231]. To investigate whether *lcla1* was affected for photosynthetic growth, tenfold dilution series was performed on minimal media and cultured in conditions of increasing light intensity (from ~2.5 - ~300 μmol photons m^{-2} s^{-1}). Interestingly, *lcla1* grew optimally on minimal media in low light conditions (~30 μmol photons m^{-2} s^{-1}) and was severely affected for growth in the highest light condition (Figure 35B) compared to control strains.

Chlamydomonas photosynthetic mutants can often be rescued in low light with supplemental acetate [230]. To investigate whether growth of *lcla1* could be restored in these conditions, tenfold dilution series was performed on acetate containing media in increasing light intensity (from 0 - ~300 μmol photons m^{-2} s^{-1}). Interestingly, growth of *lcla1* on acetate containing media mimicked that of minimal media for each light intensity

suggesting that the growth defect observed on minimal media may be due to light sensitivity (Figure 35B and 41). Indeed, the difference in growth for *lcla1*, [arg+] (or EV-transformed) versus [arg-] haploid progeny became more distinguished in the highest light condition, with *lcla1*, [arg-] haploid progeny appearing photobleached after ~5 days on media containing acetate (Figure 35B and 41).

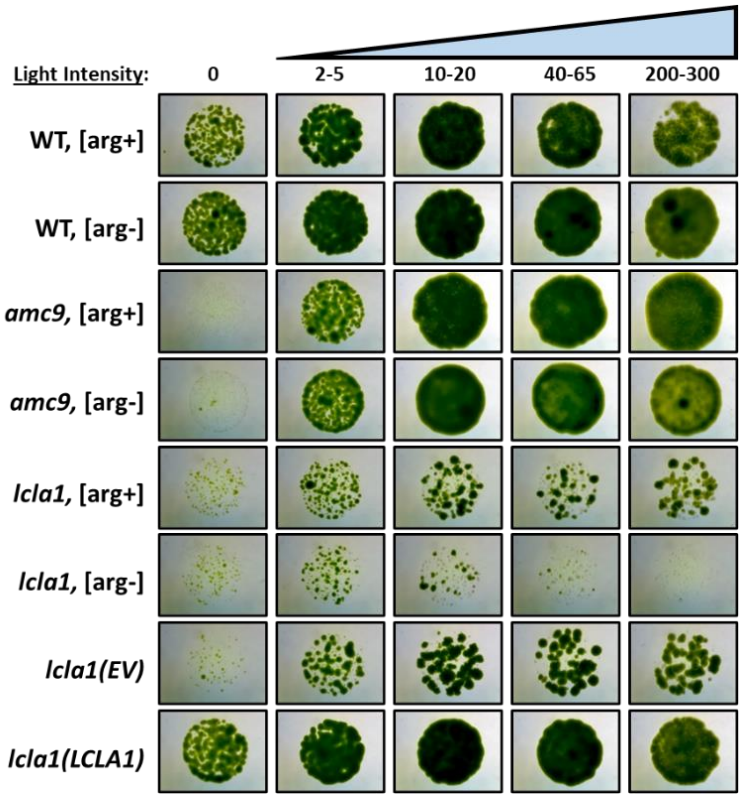


Figure 41 *lcla1*, [arg-] spores are sensitive to high light.

Magnified images of 10^{-2} dilution from tenfold dilution series of the indicated strains performed in mixotrophic conditions with increasing light intensity (µmol photons $m^{-2}s^{-1}$). Pictures of each dilution spotted all from same plate after 10 days incubation to observe photobleaching of *lcl1* in high light. Representative images chosen of WT, *amc9*, and *lcla1* spores. *lcla1* [arg+] and [arg-] spores. Transformed strains (with parentheses) carry wild-type *ARG7* gene and therefore are [arg+].

Subunits physically interacting within a larger complex, including components of the photosynthetic electron transport chain, often depend on each other for stability and loss of one can result in the proteolytic degradation of the other subunit(s) [232-234]. Therefore, assaying the abundance of one subunit can be used as a proxy for the overall complex. To investigate whether the defect in photosynthetic growth and light sensitivity was due to a defect in the protein abundance of the photosynthetic complexes, immunoblotting was performed on total protein extracted from mixotrophic cultures of *lcla1* compared to control strains. Interestingly, no differences were observed in the abundance of markers for the oxygen evolving complex, photosystem II, photosystem I, the b_6f complex, or plastid ATP synthase in *lcla1* compared to controls (Figure 42). Taken together, it was concluded that the photosynthetic growth defect and high light sensitivity observed in *lcla1* are not due to a defect in photosynthetic complex subunit abundances, at least in the growth conditions tested.

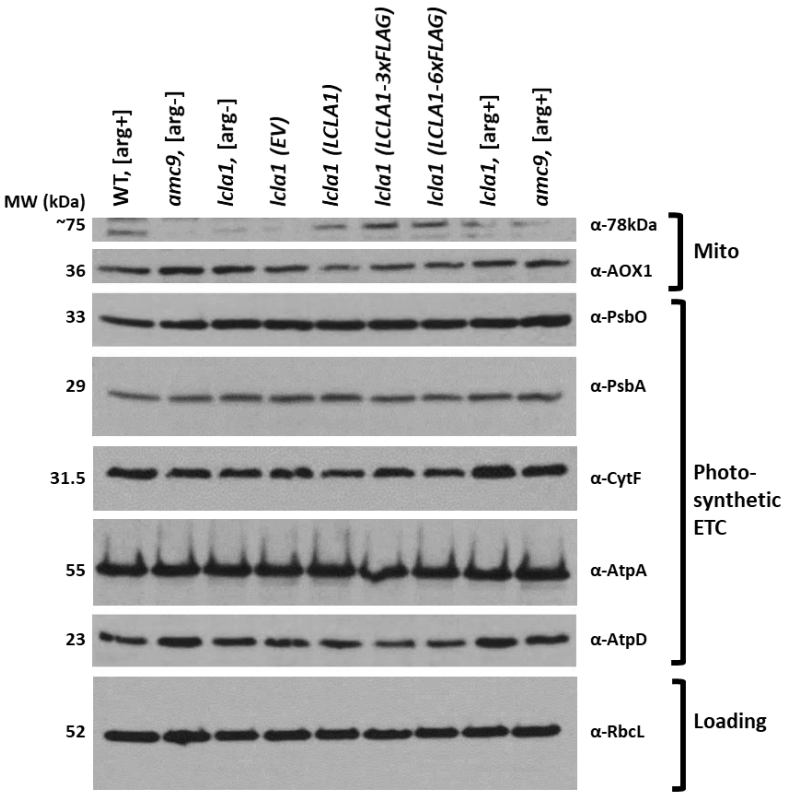


Figure 42 Chloroplast protein markers not affected in *lcla1* mutant.

Samples were collected from mixotrophic cultures shaking at 180 rpm, under constant light (50-60 µmol photons $m^{-2}s^{-1}$) with daily dilutions. Whole cell extraction was performed by freeze thaw method and total protein resolved on 12% polyacrylamide gels by SDS-PAGE (20 ug per well). Antibodies against the indicated proteins were used to diagnose potential defect in either the mitochondria (top two panels) or chloroplast (remaining panels). Arginine prototrophic strains used: WT (CC-4533), *lcla1#5*, *amc9PH39-3*, *lcla1(EV)* is #2, *lcla1(LCLA1)* is #19. Arginine auxotrophic strains used: *lcla1 spore #1*, and *amc9#PH41-27*.

4.4.5 *lcla1* phenocopies growth deficiency of the *Chlamydomonas* isocitrate lyase mutant

In *Chlamydomonas*, the isocitrate lyase [null] mutant is also affected for respiratory growth and is sensitive to high light but only in mixotrophic conditions as it is not defective for photoautotrophic growth [17]. Isocitrate lyase (ICL) catalyzes the initial step of the glyoxylate cycle, producing glyoxylate (which enters its namesake cycle) and succinate that can re-enter the TCA cycle or fuel gluconeogenesis (Figure 43A). *Chlamydomonas* adapts to loss of ICL by decreasing abundance of MAS1, CIS2, and ACH1, effectively redirecting metabolism away from the glyoxylate cycle towards the L-arginine biosynthetic pathway [17] (and summarized in Figure 43B). Since acetate assimilation is connected to the glyoxylate cycle via citrate synthase (CIS2 in *Chlamydomonas*) and is required for growth in respiratory conditions, loss of ICL is lethal in respiratory conditions [17]. The glyoxylate cycle is a biochemical bypass/shunt of the TCA cycle, as is the L-arginine biosynthetic pathway [15, 17]. So one intriguing possibility is that the observed metabolic rerouting through the arginine biosynthetic pathway is due to its ability to "complete the cycle" through the concomitant production of fumarate by arginino-succinate lyase. Alternatively, this could be an route to utilize acetyl-CoA, recycling CoenzymeA via the first step of the pathway catalyzed by glutamate N-acetyltransferase. In either case, blocking the L-arginine biosynthetic pathway by mutation would be expected to modulate mixotrophic growth – as is the case in the *lcla1* mutant. Interestingly, a severe selection bias was observed in genetic crosses to obtain *icl*, [arg-] haploid progeny, as was the case for *lcla1,* [arg-] haploid progeny (each ~9:1 [arg+]:[arg-]) but importantly not for *amc9*,

[arg-] (Figure 44A-B) or wild-type (data not shown). This bias indicated that the combination of the *icl* mutation in an *arg7* background is deleterious. As expected, *icl,* [arg-] haploid progeny are more severely affected for growth in the light whereas no change was observed for growth in respiratory conditions (Figure 44C). Furthermore, *icl,* [arg-] haploid progeny were more sensitive to canavanine supplementation in mixotrophic conditions (normal light) than *lcla1,* [arg-]. Taken together, these results underline the combined roles of the L-arginine biosynthetic pathway and glyoxylate cycle to growth of *Chlamydomonas* in mixotrophic conditions.

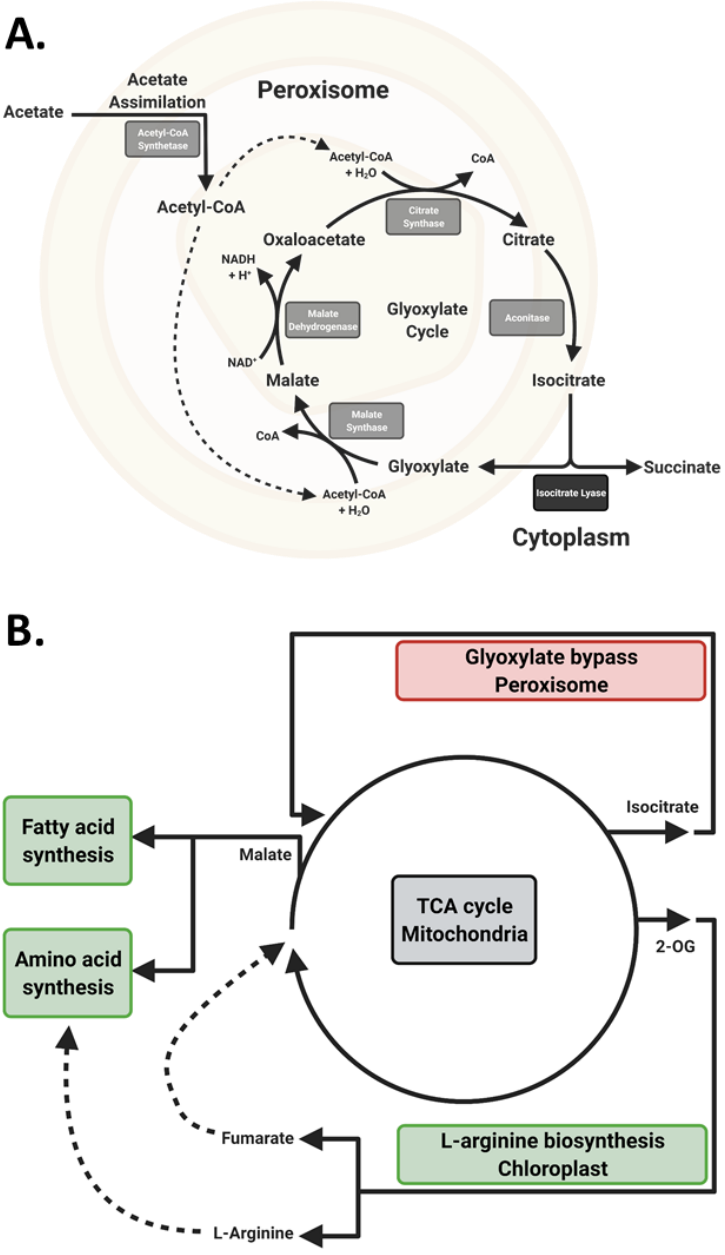


Figure 43 Model of metabolic adjustment in *Chlamydomonas icl* mutant.

A) Diagram of *Chlamydomonas* glyoxylate cycle with black box highlighting ICL and grey boxes indicating other enzymes of glyoxylate cycle (housed in the peroxisome) or participating in acetate assimilation (cytoplasm). B) Summary of metabolic response of *icl* mutant characterized by Plancke, *et al.* 2013. Green boxes indicate pathways that are increased whereas the red box representing the glyoxylate cycle indicates a decrease in

enzyme abundance. The grey box, representing enzymes of the TCA cycle were largely unaffected. Models created using BioRender.com.

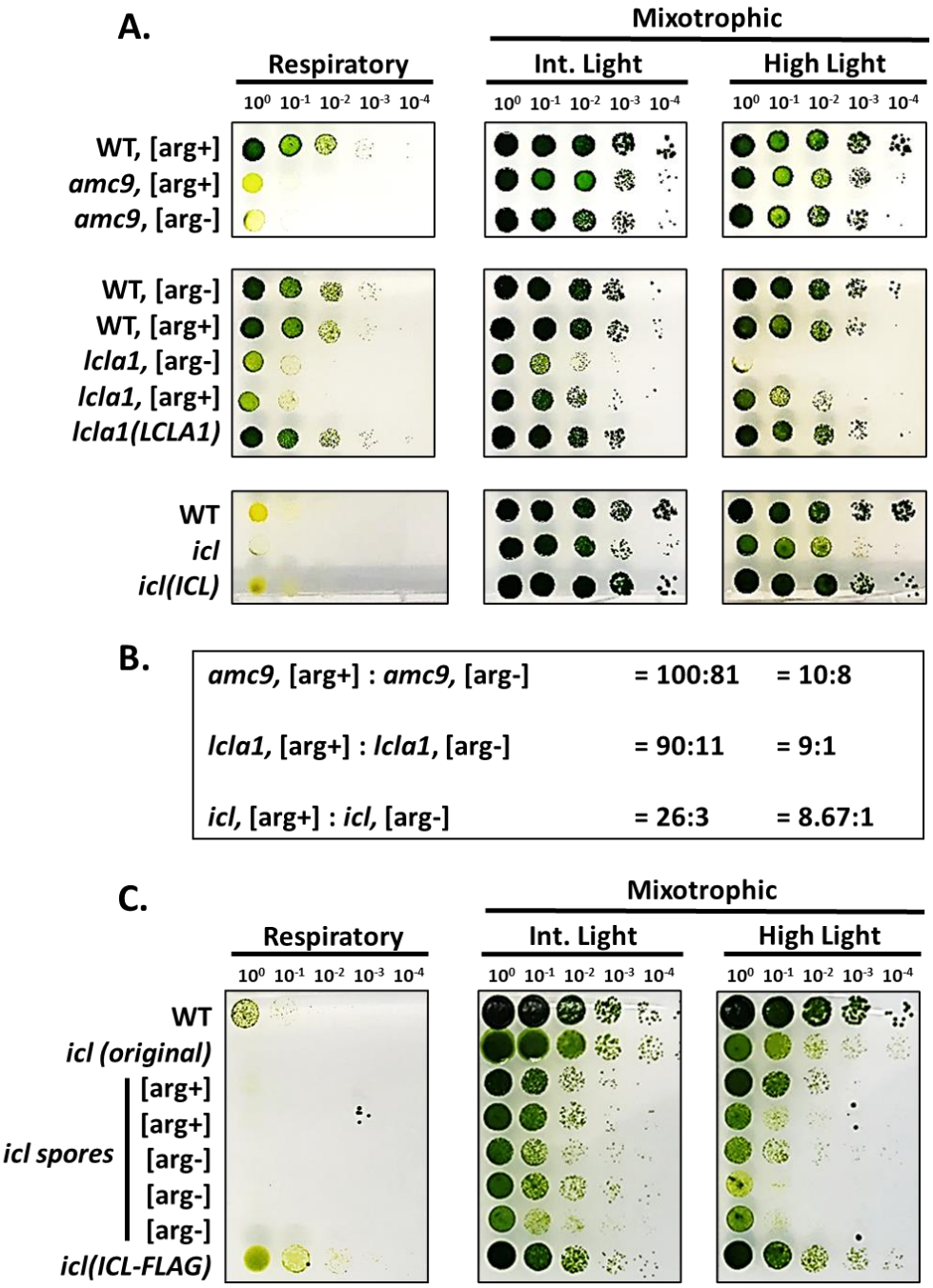


Figure 44 The *lcla1* and *icl* mutants are modulated by *arg7-8* marker.

A) Tenfold dilution series in different light intensities: dark, intermediate, or high (0, 30-50, or 200-300 µmol photons $m^{-2}s^{-1}$, respectively) for the indicated strains. A) top panel, representative *amc9*, [arg+] and [arg-] spores. A) middle panel, representative *lcla1*, [arg+]

and [arg-] spores and control strains. A) bottom panel, WT (CC-125) and *icl* mutant and complemented strain from Plancke, *et al.* 2013 shows light sensitivity of *icl* mutant in high light (all strains in this panel are [arg+]). B) Clear selection bias observed against *icl*, [arg-] spores and *lcla1*, [arg-] spores compared to [arg+] from sexual crosses. C) Tenfold dilution series in different light intensities: dark, intermediate, or high (0, 30-50, or 200-300 µmol photons $m^{-2}s^{-1}$, respectively) for the indicated strains. *icl* spores resulting from cross of *icl(ICL-FLAG)* x WT, [arg-].

4.4.6 *lcla1* and *icl* have increased TAG in non-stress conditions

In mixotrophic growth conditions the *icl* mutant was found to have increased free fatty acid content (C-18 chain length) as determined by GC-MS [17]. *Chlamydomonas* naturally stores excess carbon in the form of triacylglycerols (TAGs) under stress conditions or in the presence of mutations that affect metabolism [7, 194, 195, 209]. To investigate whether *lcla1* also had increased fatty acids in non-stress conditions, total lipids were extracted from the *lcla1* mutant and control strains then resolved by HPLC. Peaks corresponding to each TAG species were quantified, normalized for dry weight, and graphed relative to wild-type abundance. For proof of principle, the abundance of TAGs in nitrogen deprived samples (a well-established stress condition for induction of TAGs) was increased in these experiments compared to the nitrogen replete conditions (data not shown) [7, 194, 195]. As an additional control, TAGs content for the *icl* mutant compared to its respective wild-type and complemented strain were measured in the same growth conditions (Figure 45A), corresponding to that described previously. Two independent haploid progeny of *lcla1* as well as an EV-transformed strain had ~2-fold increase in TAGs compared to wild-type and complemented strains (Figure 45B).

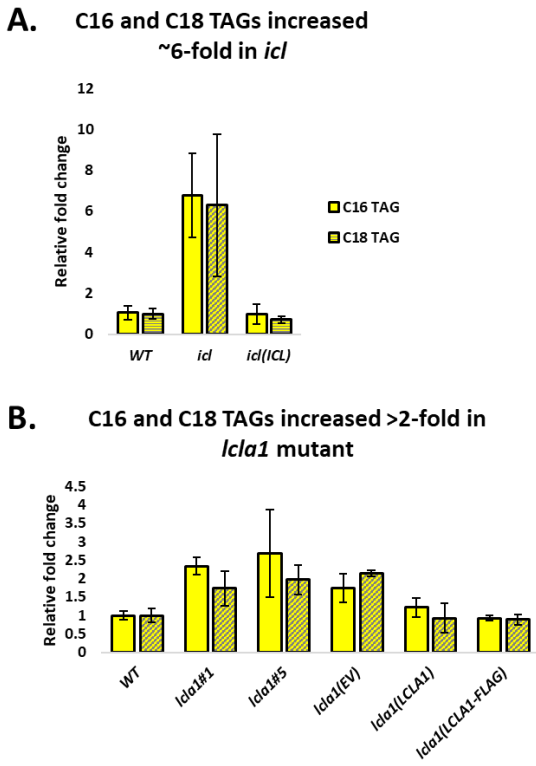


Figure 45 *icl* and *lcla1* mutants have increased TAGs.

Total lipids were extracted from mixotrophic, exponential phase cultures and resolved by HPLC. A) C18 and C16 length TAGs were identified based on comparison to C17 and C19 TAG standards (see materials and methods) and quantified for the *icl* mutant compared to WT (CC-125) and complemented strain, *icl(ICL)*. B) Same as in (A) but for *lcla1* spores, empty vector transformant, compared to WT (CC-4533) and *lcla1* complemented strains. *lcla1#1* is [arg-], *lcla1#5* and all other strains are [arg+].

TAGs are stored in specialized, single-membrane organelles called lipid bodies or lipid droplets [197, 235, 236] and can be detected by staining with lipophilic dyes such as Nile Red or BODIPY probes [7, 197, 236, 237]. To confirm the increase in TAGs in the *lcla1* mutant, lipid bodies were visualized by staining with Nile Red and subsequent confocal fluorescent microscopy (in the same culture conditions in which increased TAGs were observed previously). In agreement with the HPLC analysis, the *lcla1* recipient strain and empty vector transformant had ~two-fold increase in the average number of Nile Red staining foci per cell (Figure 46A-B). Furthermore, average size of lipid bodies, deduced from total fluorescence per lipid body were larger compared to control strains (Figure 46A, C). From these combined results we concluded that TAGs are increased in the *lcla1* mutant under normal growth conditions.

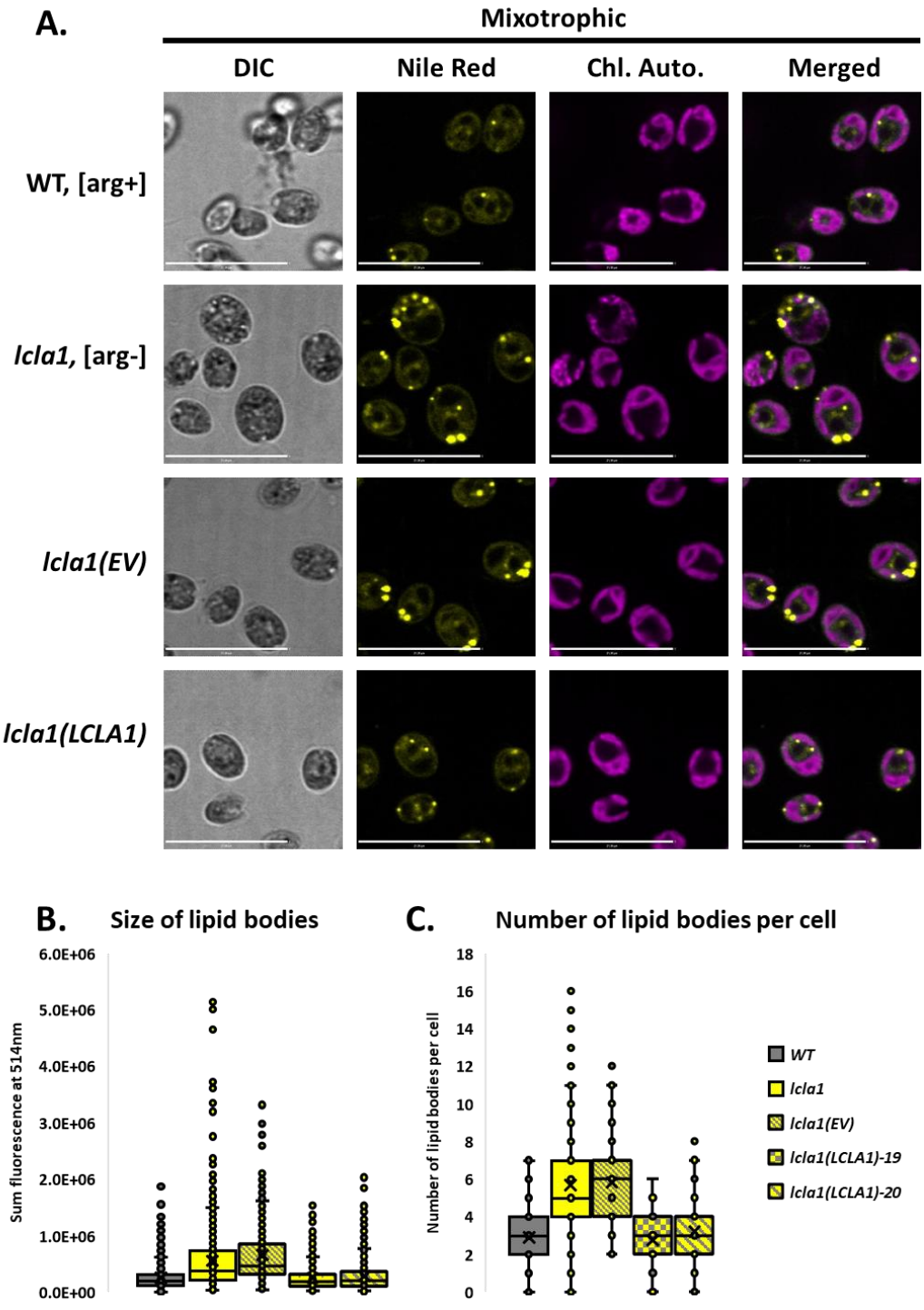


Figure 46 Lipid bodies accumulate in the *lcla1* mutant.

A) Representative images of Nile Red stained *Chlamydomonas* strains cultured in mixotrophic conditions (exponential phase, 3-5 biological replicates per strain). Images are

600x total magnification, 175x175 pixels each. Scale bar is 24 µm. B) Size of lipid bodies was inferred from total fluorescence excited at 514 nm, quantified using Volocity software and displayed by box and whisker plot. WT is CC-4533 (1,035 total lipid bodies counted), *lcla1, arg-* is *lcla1#1* (1,063 total lipid bodies counted), *lcla1(EV)#2* (655 total lipid bodies counted), *lcla1(LCLA1)#19* (643 total lipid bodies counted), and *lcla1(LCLA1)#20* (838 total lipid bodies counted). C) Lipid bodies were counted manually, (150-250 total cells counted per sample). Variation in (B) and (C) displayed as box and whisker plot.

4.4.7 High light sensitivity of *lcla1* due to lipotoxicity

High light is a stress condition known to induce fatty acid synthesis in addition to nitrogen or metal (such as iron) deprivation, high salt, *etc.* (reviewed in [236]). In *Chlamydomonas*, acetyl-CoA is used to synthesize fatty acids via a plastid-localized fatty acid synthetase [238, 239] (Figure 47A). Previously, treatment of wild-type *Chlamydomonas* with the fatty acid synthesis inhibitor cerulenin blocked accumulation of lipid bodies in nitrogen-deplete conditions [239]. Since *lcla1* had a higher basal level of lipids in the form of TAGs in mixotrophic conditions, we asked whether the high light sensitivity could be due to a toxic buildup of lipids or lipotoxicity. To address this question, tenfold dilution series was performed on TAP media containing increasing concentrations of cerulenin in normal light, high light, or in the dark. Interestingly, the effect of cerulenin was more toxic with increasing light intensity (Figure 47B-C). In mixotrophic conditions, *lcla1* haploid progeny grew comparable to wild-type and complemented strains when treated with cerulenin (Figure 47B). However, in high light stress conditions, *lcla1* haploid progeny were resistant to pharmacological inhibition of fatty acid synthesis (Figure 47B-C). No effect on growth was observed in the dark between cerulenin treated and untreated (data not shown). C75, another inhibitor for fatty acid synthesis with similar chemical structure and properties to cerulenin was also evaluated, however no effect on growth was observed when supplemented to the media even at four-fold the effective concentration of cerulenin in the same growth conditions (data not shown). It is important to note that some variation of lipotoxicity was observed among tetrads (Figure 47C and not shown)

presumably due to variation in cell wall robustness. Indeed, it has been observed that wild-type strains with more robust cell wall are more resistant to different antibiotics (data not shown).

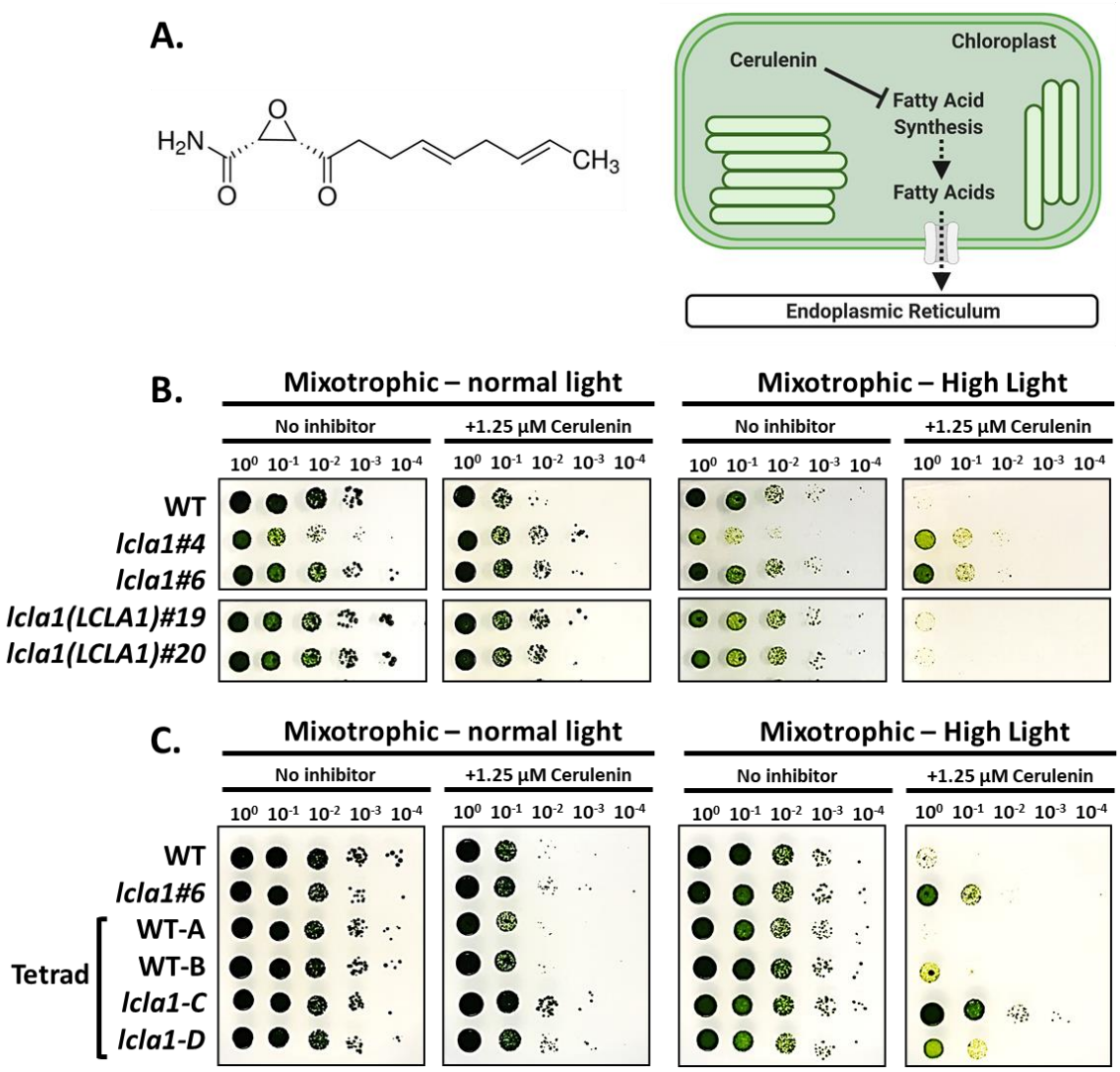


Figure 47 *lcla1* spores are resistant to FAS inhibition.

A) Model for inhibition of plastid fatty acid synthase by cerulenin created with BioRender.com (chemical structure courtesy of Sigma-Aldrich). B and C) Tenfold dilution series on media containing 1.25 µM cerulenin compared to no drug in the indicated conditions (30 µmol photons $m^{-2}s^{-1}$ for normal light, 200-300 µmol photons $m^{-2}s^{-1}$ for high light). B) Comparison of sensitivity for WT (CC-4533) and complemented strains versus resistance of *lcla1#4* [arg-] and *lcla1#6* [arg+] spores. C) Segregation of cerulenin resistance in SID, Hyg^R spores of sample tetrad compared to sensitivity of WT parental strain (CC-4533). No affect of cerulenin on growth was observed in respiratory conditions (not shown). All images of plates taken after 7 days incubation in indicated conditions.

Conceptually, lipids can accumulate due to increased lipid synthesis or decreased lipid catabolism. Since *lcla1* had ~two-fold increase in TAGs in mixotrophic conditions, we then asked whether it was able to utilize these lipid reserves to support growth in conditions where any other carbon source is absent. To address this question, *lcla1* and wild-type cells were cultured in mixotrophic conditions (normal light intensity), washed, and resuspended in minimal media (with no carbon source) and cultivated in the dark. Cell density was measured twice daily and plotted as a function of time. Wild-type cells continued to divide for ~24 hours post-inoculation after which cell density decreased over the following ~24 hours. In stark contrast, the cell density for *lcla1* decreased dramatically post-inoculation and continued to decline steadily for the next two readings (Figure 48A). In contrast to mammalian model systems, catabolism of lipids via fatty acid oxidation (β-oxidation) in plants and fungi occurs exclusively in the peroxisome [235]. It has previously been shown that yeast mutants defective in peroxisome function are sensitive to oleic acid since it cannot be catabolized [200]. Similarly, in *Chlamydomonas* oleic acid was unable to support growth as the sole carbon source of a mutant deficient in ACX2, a peroxisome localized acyl-CoA oxidase which produces trans-enoyl-CoA from their respective acyl-CoA substrates, a process central to β-oxidation [19]. To investigate whether an exogenously supplied fatty acid could support growth of *lcla1*, cultures were prepared as described above and oleic acid added both at inoculation and again ~24 hours later. Cell density decreased significantly for *lcla1* after inoculation and there was no further growth, analogous to samples without oleic acid supplementation described above (Figure 48B-C). In contrast, treatment with oleic acid boosted growth of the wild-type strain following

overnight incubation of each oleic acid feeding (Figure 48B-C). Taken together, these results suggest that *lcla1* is defective in lipid catabolism.

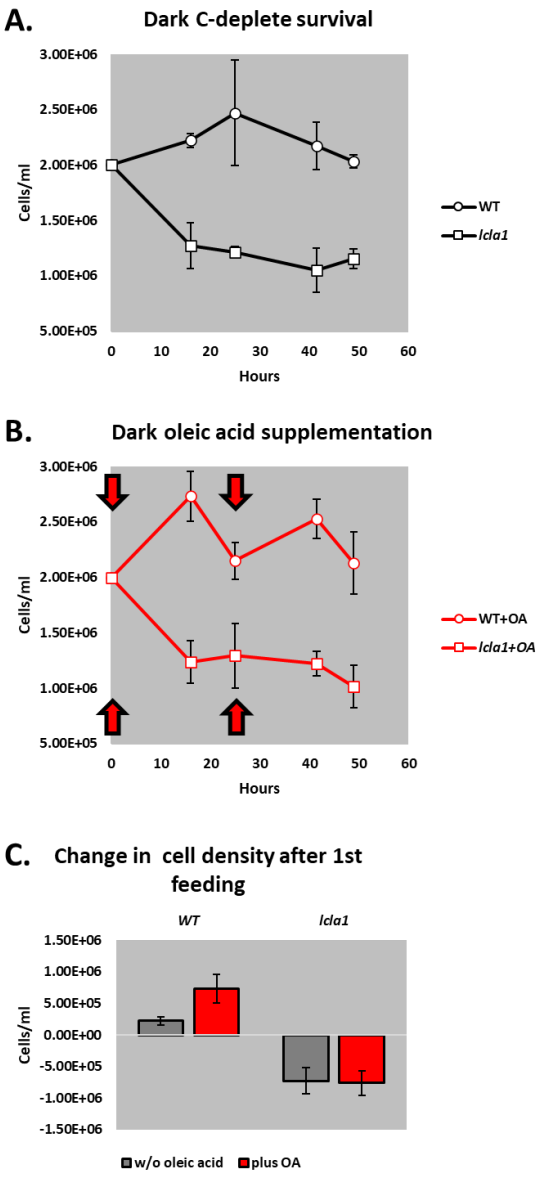


Figure 48 *lcla1* is unable to utilize stored lipids.

Mixotrophic exponential phase cultures of WT (CC-4533) and *lcla1* (*lcla1#6*) were collected, washed, and resuspended to $2x10^6$ cells/ml (mid-exponential phase) in minimal media and transferred to the dark. A and B) Cell density was measured twice daily and plotted as a function of time for three biological replicates when (A) no exogenous carbon source or (B) oleic acid (OA), an 18-carbon fatty acid, was supplemented. B) Oleic acid was added at each indicated time point (red arrows) to a final concentration of 0.5 mM. C) Change in cell density was calculated after the first oleic acid feeding and plotted for each strain. Both strains were [arg+] so no supplemental L-arginine was required.

4.4.8 N-terminus of LCLA1^S targets reporter to discrete cellular compartment

To further explore possible functions of LCLA1, identification of its subcellular localization was critical. PredAlgo, an alga specific protein localization prediction algorithm [146] predicted mitochondrial localization for both isoform. However, PredAlgo is notably unreliable so several additional localization prediction tools were queried [144, 145, 147, 210-212]. Unfortunately, there was no consensus in prediction results for LCLA1^S or LCLA1^L, or even for several *Chlamydomonas* proteins of known compartmentalization (Table 6). As stated above, β-oxidation of fatty acids occurs exclusively in peroxisomes in plants and the observed defects in fatty acid metabolism in the *lcla1* mutant led us to hypothesized peroxisomal function was affected due to loss of LCLA1. In such scenario, it is possible the short and/or the long LCLA1 isoform localizes to the peroxisome. Manual inspection of LCLA1^L amino acid sequence compared to peroxisome resident enzymes identified a putative PTS2 (Type II Peroxisomal Targeting Sequence at the N-terminus (RL-X_4-QLL) very similar to that of *Chlamydomonas* ASC2 (RL-X_5-QL) [15]. Previously, the N-terminal 25 amino acids of ACS2, including this motif, was found sufficient to target a fluorescent reporter to peroxisomes in *Chlamydomonas* [15]. As described above, the LCLA1^L isoform is so-named due to an N-terminal extension which is absent in LCLA1^S. Therefore, LCLA1^S lacks the putative PTS2 of LCLA1^L. Import of peroxisomal proteins is under the control of several pathways [240]. One pathway controls the import of substrates with the canonical PTS1 tripeptide (Type I Peroxisomal Targeting Signal, for example, SKL*) at the very C-terminus of proteins

[240]. The SKL tripeptide is present in LCLA but is 381 amino acids from the C-terminus (in each isoform) and therefore unlikely to be a true PTS1 although an example of an internal PTS1 signal has been found in plants previously [241]. Aside from targeting via the presence of a PTS1, another route uses a PTS2 motif at the N-terminus of protein substrates. A third route, less characterized, utilizes a PTS3 (Type III Peroxisomal Targeting Signal) signal patch [242]. Other import mechanisms such as by way of another organelle such as the ER or mitochondria also have been described [240, 243]. Subsequent sequence analysis focusing on the N-terminus of LCLA1^{S} revealed the presence of a putative membrane peroxisomal targeting signal (mPTS, Figure 49A) for peroxisomal membrane proteins (PMP), characterized by a cluster of basic residues within a predicted α-helix [244].

Gene	Protein	Experimental	DeeplocP	mitofates	PredAlgo	Targetp 1.1	targetP 2.0	SherLoc2	WoLF PSORT
Cre07.g329861	Creo7.g329861.t1.2	*n.d.*	peroxisome, soluble	no mito presequence	Mitochondria	mitochondria	other	cytoplasmic	chloroplast
Cre07.g329861	Creo7.g329861.t2.1	*n.d.*	peroxisome, soluble	no mito presequence	Mitochondria	mitochondria	other	mitochondrial	ER
Cre07.g329861	Creo7.g329861.t2.2	*n.d.*	peroxisome, soluble	no mito presequence	Mitochondria	mitochondria	other	mitochondrial	chloroplast
Cre07.g329861	Cre07.g329861.kozak.1	*n.d.*	peroxisome, soluble	no mito presequence	Secretory Pathway	other	other	extracellular	chloroplast
Cre07.g329861	Cre07.g329861.kozak.2	*n.d.*	peroxisome, soluble	no mito presequence	other	other	other	cytoplasmic	ER
KXZ53939.1	Putative Gonium ortholog	*n.d.*	peroxisome, soluble	Mito presequence	Secretory Pathway	mitochondria	other	extracellular	plasma membrane
XP_002950855.1	Putative Volvox ortholog	*n.d.*	nucleus, soluble	no mito presequence	Secretory Pathway	other	other	nuclear	chloroplast
Cre09.g415500	TCA1	Chloroplast	nucleus, membrane	no mito presequence	Mitochondria	mitochondria	other	mitochondrial	chloroplast
Cre08.g358250	MCA1	Chloroplast	chloroplast, memrane	no mito presequence	Chloroplast	chloroplast	other	chloroplast	chloroplast
Cre17.g700950	FDX5	Chloroplast	mitochondria, soluble	Mito presequence	Chloroplast	mitochondria	chloroplast transit peptide	mitochondrial	chloroplast
Cre13.g603700	ACTIN	cytoplasm	cytoplasm, soluble	no mito presequence	other	other	other	cytoplasmic	cytoskeleton
Cre03.g181600	UPF0565 PROTEIN C2ORF69	cytoplasm	cytoplasm, soluble	no mito presequence	other	other	other	cytoplasmic	cytoplasm
Cre06.g282800	ICL1	cytoplasm	peroxisome, soluble	no mito presequence	other	other	other	peroxisomal	mitochondria
Cre17.g738350	CrLPAAT2	ER	ER, membrane	no mito presequence	Secretory Pathway	secretory pathway	other	vacuolar	plasma membrane
Cre08.g371052	IRE1	ER	nucleus, soluble	no mito presequence	Mitochondria	other	other	cytoplasmic	plasma membrane
Cre07.g339900	SERINE-THREONINE PROTEIN KINASE	Golgi + secretory path	cell membrane, membrane	no mito presequence	Secretory Pathway	secretory pathway	Signal peptide	plasma membrane	plasma membrane
Cre10.g457750	SLC26, SULTR4	Golgi + secretory path	cell membrane, membrane	no mito presequence	other	other	other	cytoplasmic	plasma membrane
Cre12.g483950	MDH4	Mitochondrial	mitochondria, soluble	no mito presequence	Chloroplast	chloroplast	chloroplast transit peptide	cytoplasmic	chloroplast
Cre12.g542500	MOC1	Mitochondrial	mitochondria, soluble	no mito presequence	Secretory Pathway	other	other	cytoplasmic	nuclear
Cre16.g688900	AMC1	Mitochondrial	mitochondria, soluble	Mito presequence	Mitochondria	mitochondria	other	mitochondrial	chloroplast
Cre12.g507300	LCI30	nucleus	nucleus, soluble	no mito presequence	other	other	other	nuclear	nuclear
Cre06.g274200	Histone H2A	nucleus	nucleus, soluble	no mito presequence	Chloroplast	other	other	nuclear	nuclear
Cre09.g390023	CRR1	nucleus	nucleus, soluble	no mito presequence	other	other	other	nuclear	nuclear
Cre10.g423250	MDH2	peroxisome	mitochondria, soluble	no mito presequence	other	other	other	cytoplasmic	cytoplasm

Table 6 Subcellular localization prediction of LCLA1 isoforms.

Subcellular localization was predicted using 7 different prediction algorithms for possible LCLA1 isoforms (grey background) compared to proteins of known compartmentalization in *Chlamydomonas*. Correct predictions are highlighted in green. Light green is used for similarly targeted compartments.

Subcellular localization is often facilitated through sequence motifs present within the N-terminal region of a nascent protein. Since putative motifs were found in the N-terminus of $LCLA1^{L}$ and $LCLA1^{S}$, we adopted a strategy in which the N-terminal region was fused to fluorescent reporter and C-terminal epitope tag. The sequence encoding the N-terminal ~70-100 amino acids of each isoform where cloned in frame of the *Chlamydomonas* codon optimized YFP, *CrVENUS*, and introduced into the wild-type strain by glass bead transformation. Notably, very few transformants were obtained by introduction of *$LCLA1^{L}$-CrVENUS*. Those that were obtained were clearly cytotoxic (data not shown) and further attempts to express this artificial protein have been unsuccessful.

In contrast, *$LCLA1^{S}$-CrVENUS* transformants were easily obtained and presence of transgene confirmed as described above (and in methods). Since transforming DNA integrates in the *Chlamydomonas* genome at random, to eliminate potential of artificial expression based on position of insertion, expected size of the chimeric protein was confirmed by immunoblotting using the anti-FLAG antibody (Figure 49B). Additionally, introduction of the transgene did not affect growth of the resulting transformants as assayed by tenfold dilution series (data not shown). In mixotrophic conditions, localization of *$LCLA1^{S}$-CrVENUS* was primarily similar to that of ER markers (Figure 50A and [201]). However, it is possible that subcellular localization could be affected based on growth condition. Since the *lcla1* mutant was most affected for respiratory growth, we next asked whether localization of *$LCLA1^{S}$-CrVENUS* changed in wild-type after a 24-hour dark adaptation (i.e. shift to respiratory growth). Intriguingly, the localization pattern of the

chimeric reporter changed resembling that of the peroxisome marker, CIS2 (Figure 50B and 51) [15].

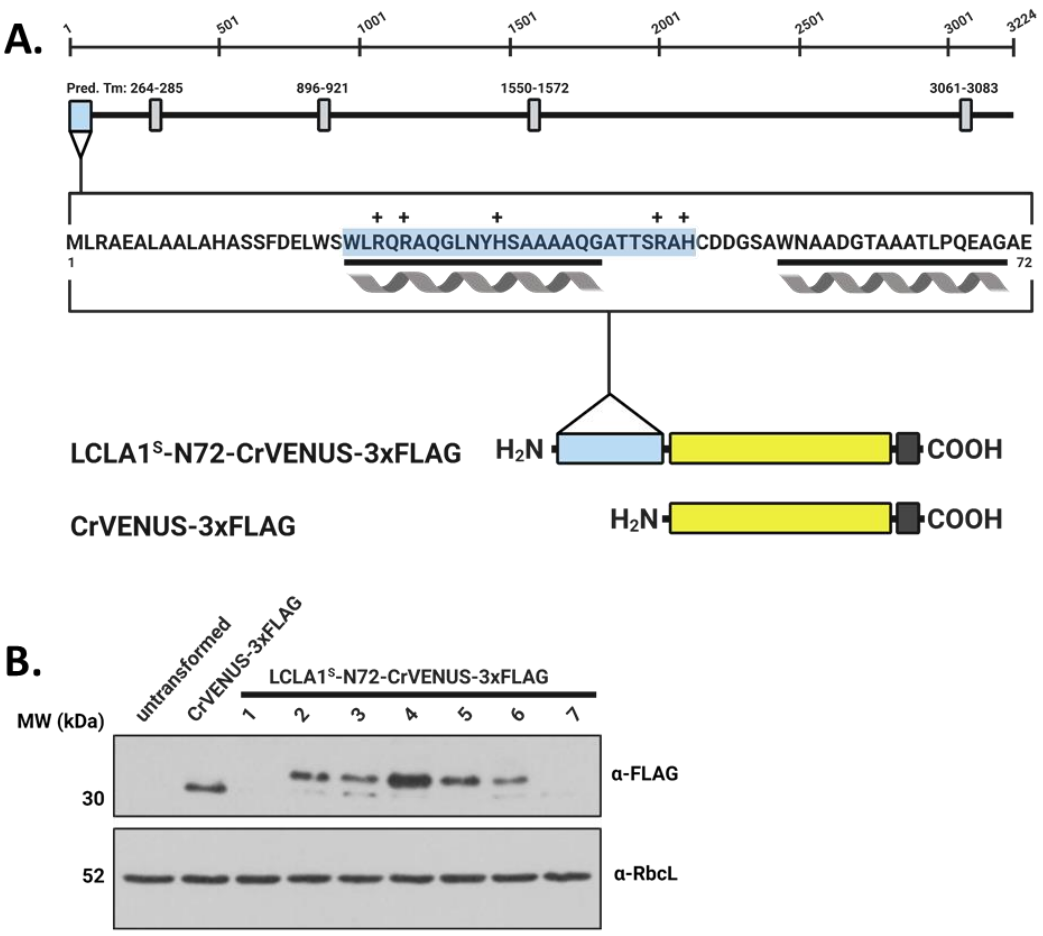


Figure 49 Model for expression of LCLA1^S chimeric fluorescent reporter.

A) Model of the LCLA1^S isoform with the 4 predicted transmembrane domains indicated by grey boxes. The primary amino acid sequence for the N-terminal 72 amino acids (blue box) is expanded with predicted α-helices designated below primary sequence with grey bars. Sequence highlighted in light blue is basic region (Trp22-His47, 26 total amino acids, has pI of 12.01, ~2x that of full sequence 6.26). The sequence encoding this region was cloned into the *Hpa*I restriction site of pMO449 in-frame of the CrVENUS fluorescent protein and designated as pADC9. B) WT (CC-4533) pADC9-transformants (encoding LCLA1^S-N72-CrVENUS-3xFLAG) were pre-screened by genotyping PCRs (not shown) and positive clones assayed for expression of the recombinant protein via immunoblot (using the α-FLAG antibody to detect the C-terminal epitope tag). Twenty µg protein from whole cell extracts was loaded per well; α-RbcL was used to assess even loading.

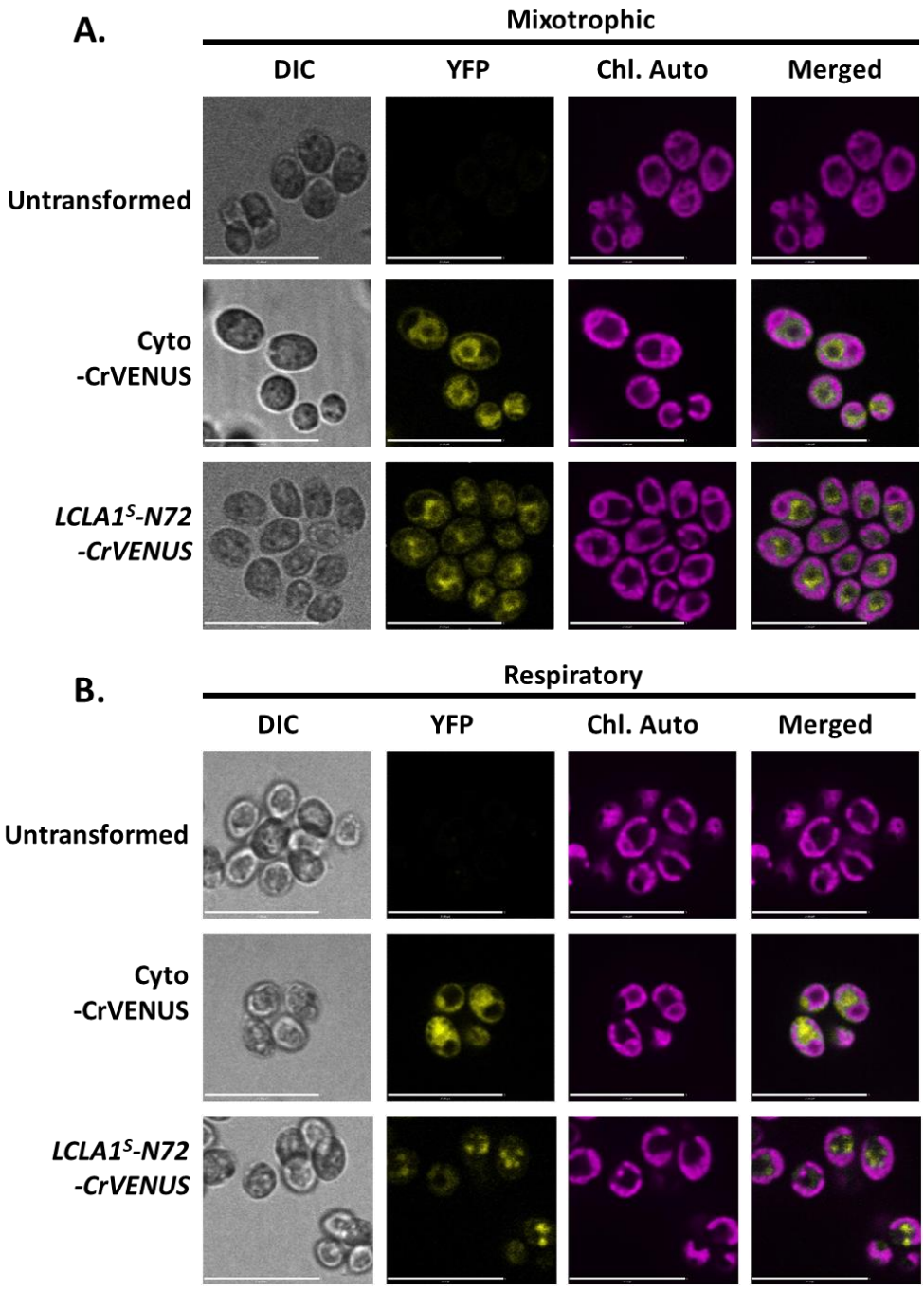


Figure 50 Subcellular localization of LCLA1^{S}-N72-CrVENUS.

Representative images of the indicated strains obtained by confocal fluorescence microscopy. Untransformed recipient is the WT strain (CC-4533). Localization of LCLA1^{S} -N72-CrVENUS compared to empty vector (pMO449) transformant expressing cytoplasmic CrVENUS in A) mixotrophic and B) respiratory conditions (after 24-hour dark

adaptation). 514nm was used for YFP excitation and 640 nm used for excitation of chlorophyll autofluorescence. Images are at 600x total magnification (150x150 pixels each), scale bar is 24 µm.

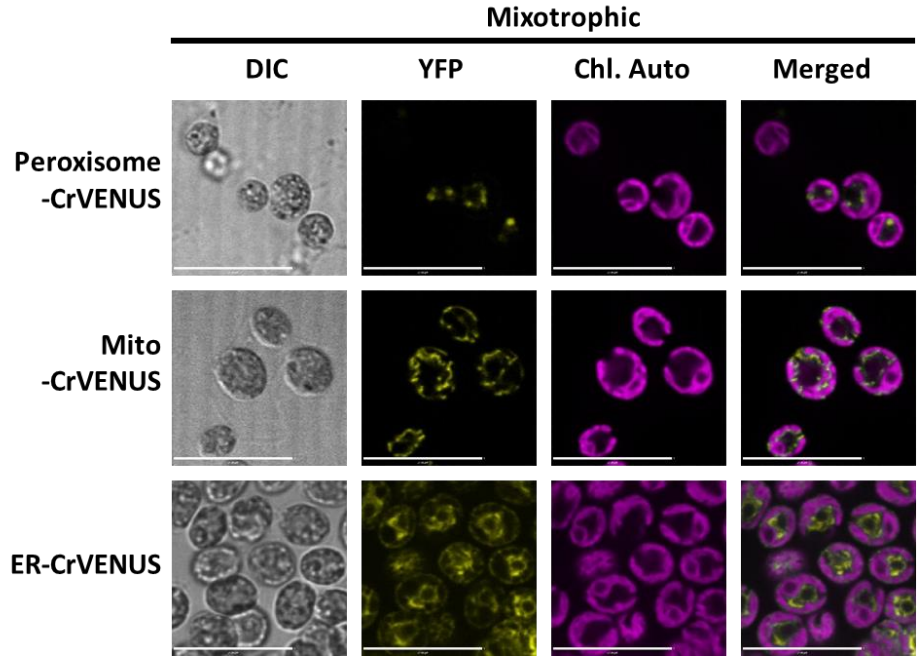


Figure 51 Subcellular localization of known organelle fluorescent markers.

Representative images of the indicated strains obtained by confocal fluorescence microscopy. The peroxisome-CrVENUS containing (CIS2 N-terminal fusion) was generously provided by Kyle Lauersen [15]. Mito-CrVENUS was created by Luke Mackinder and shown targeted to mitochondria via MDH4 N-terminal fusion [201]. YFP and chlorophyll were excited at 514nm and 640 nm respectively. All images are at 600x total magnification (150x150 pixels each), scale bar is 24 µm taken from mid-exponential phase cultures in mixotrophic conditions.

4.4.9 LCLA1 is a large low-complexity protein

As described above, *Cre07.g329861* is predicted to encode two large proteins of 3,347 and 3,224 amino acids ($LCLA1^{L}$ and $LCLA1^{S}$, respectively). However, to our knowledge neither isoform has been identified in prior *Chlamydomonas* proteomics experiments. To verify that this locus was indeed protein coding and provide further support of the revised gene model, we generated versions of the *LCLA1* gene carrying sequences corresponding to 3xFLAG and 6xFLAG at the C-terminus (see materials and methods) (Figure 32). The LCLA1-3xFLAG and 6x-FLAG were introduced into the *lcla1* mutant and the resulting transformants, assayed by tenfold dilution series showed that expression of both versions could complement the SID phenotype and high light sensitivity (Figures 52 and 53).

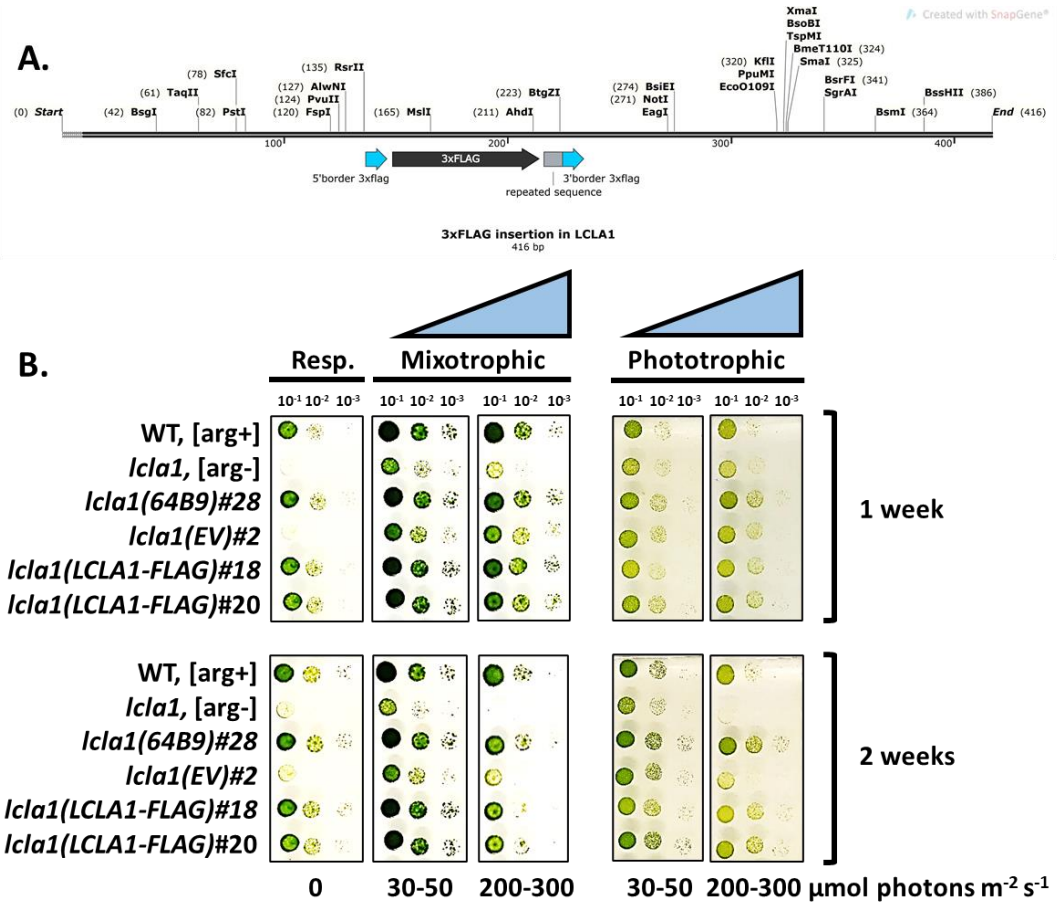


Figure 52 LCLA1-3xFLAG complements the *lcla1* mutant.

A) Map of 3xFLAG insertion in *LCLA1* of pADC10 vector. B) Tenfold dilution series of the indicated strains. WT is CC-4533, *lcla1*, [arg-] is l*cla1#1* (recipient strain for pADC10 transformation). Images for top panel taken at 1 week whereas bottom panel taken at 2 weeks. Light intensity is increasing from left to right with range (µmol photons $m^{-2}s^{-1}$) described at bottom. Minimal media is same as for TARG except lacking acetate (recipe details in methods). Transformed strains (with parentheses) are all [arg+] due to *ARG7* selectable marker of cosmid (see Table 5).

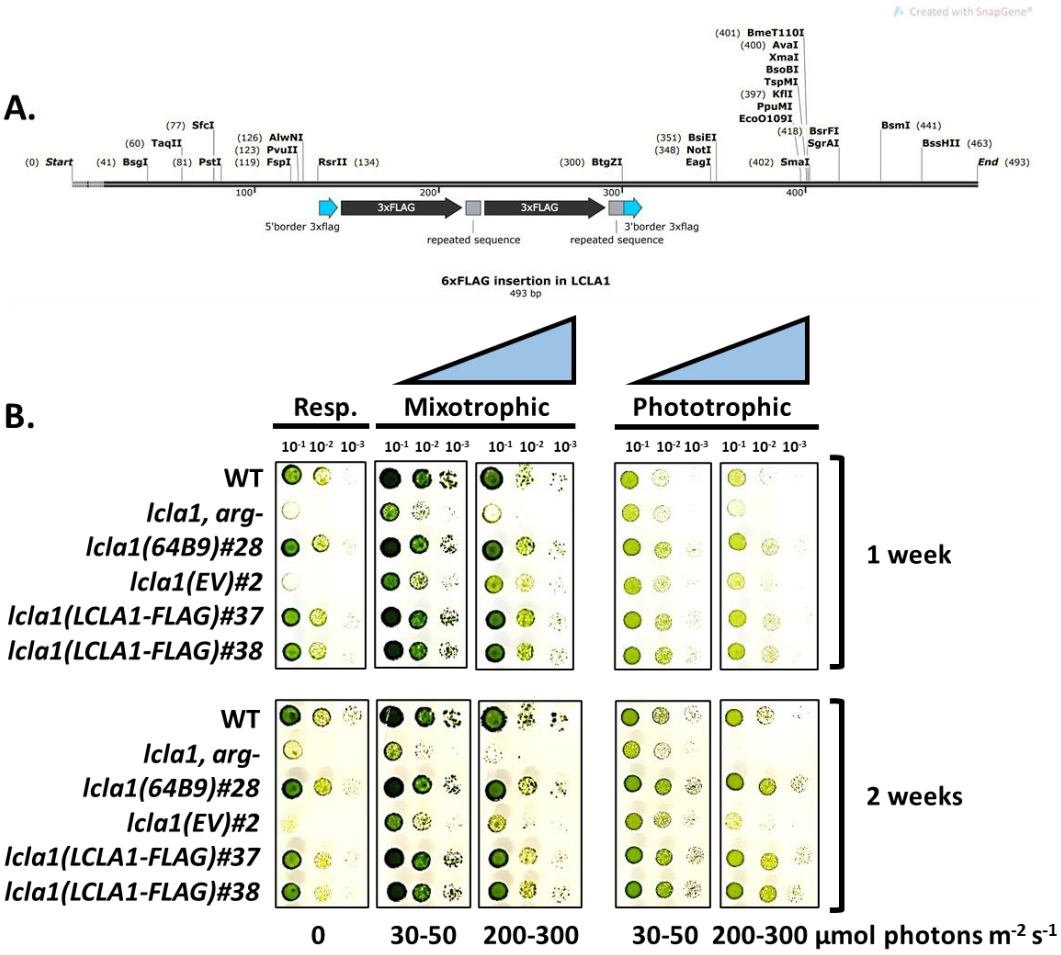


Figure 53 LCLA1-6xFLAG complements the *lcla1* mutant.

A) Map of 6xFLAG insertion in LCLA1 of pADC11 vector. B) Tenfold dilution series of the indicated strains. WT is CC-4533, *lcla1*, [arg-] is l*cla1#1* (recipient strain for pADC11 transformation). Images for top panel taken at 1 week whereas bottom panel taken at 2 weeks. Light intensity is increasing from left to right with range (µmol photons $m^{-2}s^{-1}$) described at bottom. Minimal media is same as for TARG except lacking acetate (recipe details in methods). Transformed strains (with parentheses) are all [arg+] due to *ARG7* selectable marker of cosmid (see Table 5).

Since the growth of *lcla1* was most affected in the dark and in high light, we reasoned that expression of LCLA1 was highest in these conditions. To test this hypothesis and confirm the predicted size of LCLA1, whole cell extracts from the indicated strains were collected from cultures in different growth conditions, resolved in denaturing conditions, and the recombinant protein detected by immunoblot. In each of the conditions, a single band of approximately the expected size (325 kDa) was observed (Figure 54A). However, at least two additional bands were detected migrating as lower molecular weight (~200-250 kDa). These additional bands were least prominent in mixotrophic conditions. Notably, these additional bands were specific to LCLA1 and were observed in 4 independent transformants but not the corresponding wild-type, *lcla1* recipient, empty vector transformed, or other complemented strains (Figure 54B). As expected, LCLA1-FLAG was most abundant in high light and low light but in normal light in mixotrophic conditions barely detectable (Figure 54A). In summary, it was concluded that insertion of the 3xFLAG epitope tag (or 6xFLAG) did not disrupt a domain of LCLA1 required for its function and *Cre07.g329861* is a protein coding gene. Further work is in progress to discern the identity of the faster migrating molecular weight bands.

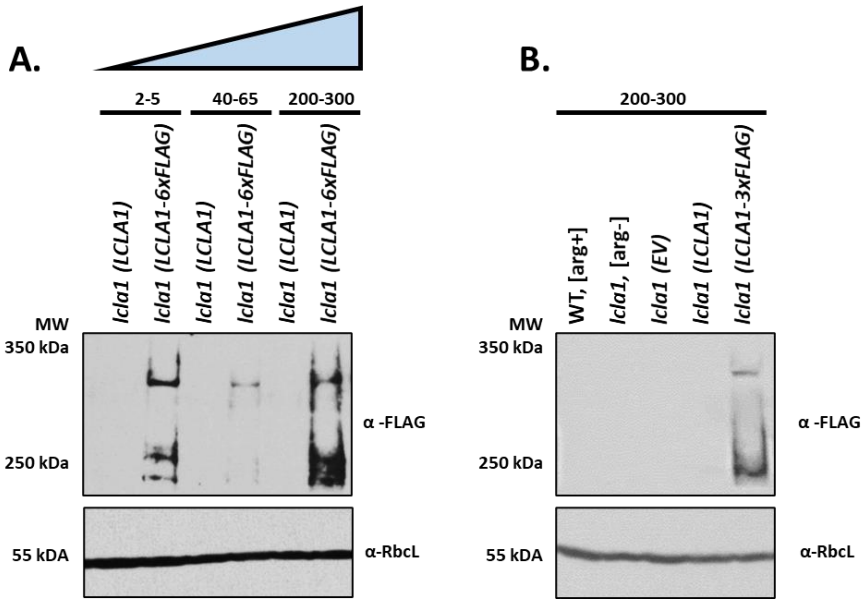


Figure 54 Multiple LCLA1 isoforms present in LCLA1-3xFLAG and -6xFLAG complemented strains.

A and B) Strains to be tested were refreshed 3x over two weeks in the indicated light conditions prior to collection. The sole exception was *lcla1#1*, [arg-], which is not viable in high light, was instead propagated in lower light, 30 µmol photons $m^{-2}s^{-1}$. Protein from whole cell lysates were then resolved in denaturing conditions on 4-10% polyacrylamide gradient gels and the recombinant LCLA1 detected by the α-FLAG antibody. α-RbcL was used to assess even loading. A) Representative immunoblot of pADC2-transformed (without FLAG tag) versus pADC11-transformed (6xFLAG) strains cultured in increasing light intensity (µmol photons $m^{-2}s^{-1}$, 150 µg per well, 3 replicates). B) No proteins of similar molecular weight cross-reacting with the α-FLAG antibody are present in control strains in high light (µmol photons $m^{-2}s^{-1}$, 160 µg per well, 4 replicates). Transformed strains (with parentheses) are all [arg+] due to *ARG7* selectable marker of cosmid (see Table 5).

Similarity searches via NCBI BLASTp identified 2 orthologous proteins in closely related algae: *Gonium pectorale* (A0A150GVS2, 29.8% identity, 36.6% similarity vs. $LCLA1^{L}$) and *Volvox carteri* (D8TWL0, 30.1% identity, 38.1% similarity vs. $LCLA1^{L}$) [216] (Figure 55). Local alignments via the Smith-Waterman algorithm were less similar but had decreased gap percentages (24.8% identity/32.2% similarity for *G. pectorale* and 23.7% identity/32.5% similarity for *V. carteri*) [216]. Additional potential orthologues were identified via BLAST searches of the Uniprot database but were also low scoring, <29.3% sequence identity with poor coverage and high percentage of sequence gaps. Results were generally the same when limiting query results to plant proteomes. Similarity searches of overlapping 500 amino acid segments of $LCLA1^{L}$ or $LCLA1^{S}$ were also performed via BLAST against Uniprot proteomes but were inconclusive.

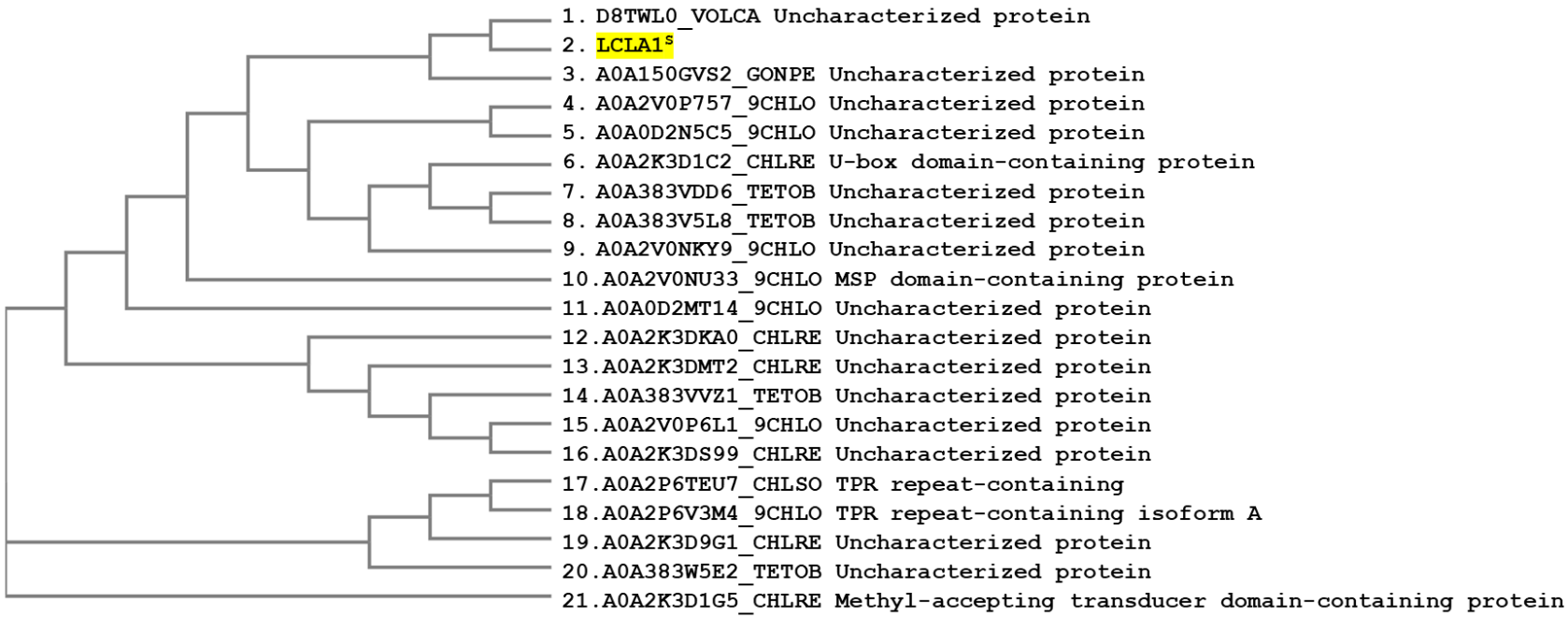
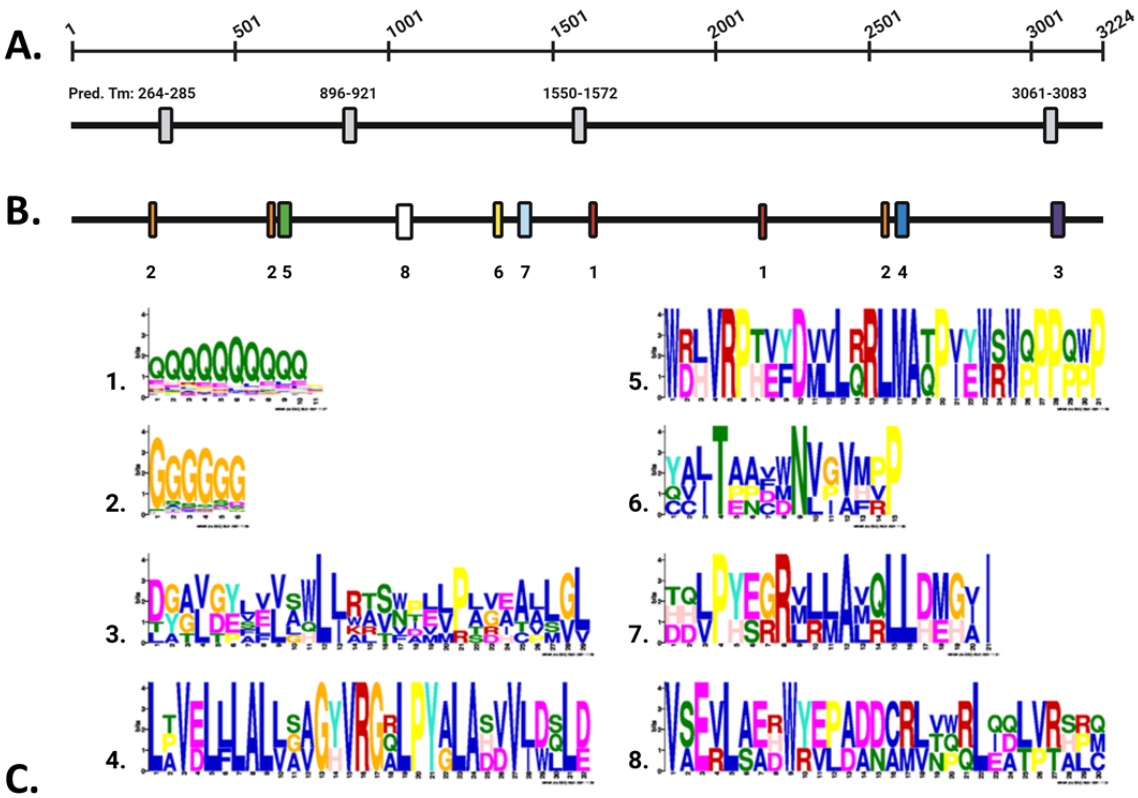


Figure 55 Evolutionary relatedness of LCLA1.

LCLA1^{S} sequence was used to identify potential orthologous proteins in plantae via the BLAST tool through the Uniprot database. The resulting top 20 BLAST hits were aligned using Clustal Omega tree-building program (EMBI) with accession number (Uniprot) given at end of each branch. Accessions in *V. carteri* (VOLCA) and *G. pectorale* (GONPE) were most similar among plants (*Tetraselmis sp. GSL018*= 9CHLO, *Chlamydomonas reinhardtii* = CHLRE, *Tetradesmus obliquus* = TETOB, *Chlorella sorokiniana* =CHLSO).

Neither LCLA1 isoform (nor putative orthologues A0A150GVS2 and D8TWL0, have any clearly conserved domains suggestive of any biological role. Using SMART, a protein sequence analysis program linked to Uniprot and Ensembl databases, many Low Complexity Regions (LCRs) were identified (covering 35.7% of LCLA1-short). LCRs are defined as sections of a primary amino acid sequence with biased composition to one or several amino acids [217, 218]. Similarly, LCRs covered 42.5% and 39.8% of the possible *G. pectorale* and *V. carteri* orthologues.

Local pairwise sequence alignments of LCLA1-long versus each putative ortholog from *G. pectorale* and *V. carteri* suggested possible evolutionary relatedness. To potentially identify any common sequence motifs conserved among the top 20 BLAST hits (limited to plant proteomes, including the putative *G. pectorale* and *V. carteri* orthologues), the MEME algorithm was used which finds short, un-gapped motifs in sequences that are manually provided with the assumption that these sequences are in fact related. [168, 191]. Eight motifs ranging from 6 to 32 amino acids in length were found (Figure 56) and submitted to FIMO, a program that queries a selected proteome for entries containing the submitted consensus sequence [168, 191]. The resulting hits (cut off at 1,000 proteins each) were then submitted to gene ontology enrichment analysis and the top hits in *A. thaliana* and *C. reinhardtii* are currently under evaluation for potential relevance.



Motif #	sequence in LCLA1	Motif length (aa)	*p*-value	Position in LCLA1^{S}	Position in LCLA1^{L}
1	QQQQQLQQSQR	11	9.77E-10	x	110-120
1	PQRQQQWQQQP	11	1.45E-09	1586-1596	1708-1718
1	EHQRRQQQLQE	11	3.40E-07	2141-2151	2263-2273
2	GGGGGP	6	3.69E+02	247-252	369-374
2	GGVGGG	6	4.54E-06	615-620	737-742
2	GGGGQG	6	5.83E-06	2537-2542	2659-2664
3	TATVGPFVVAQLLWRSVDVLPTDIAVSVL	29	7.74E-18	3072-3100	3194-3222
4	LAVELLLALVGVGYVRGQLPYALASDVLWQLE	32	1.76E-30	2646-2677	2768-2799
5	WDHVRPTVFDVVLRRLMATPIYWRWQPPQPP	31	1.27E-40	636-666	758-788
6	YALTEAFMNVGVMPP	15	7.99E-18	1289-1303	1411-1425
7	HDLPHSRRVRLAVRLLHEGII	21	9.23E-19	1393-1413	1515-1535
8	VSEVLAERWYEPAADCRLTQQLEDLVTALC	30	5.05E-29	1046-1075	1168-1197

Figure 56 MEME analysis for plant LCLA1-like proteins.

A) Model of LCLA1^{S} summarizing predicted Tm helices (manually curated). B) Location of motifs identified by MEME program numbered below to match consensus sequences. C) Summary of matching sequence, length, score, and position each motif is found in LCLA1^{S} and LCLA1^{L}. A and B) Created with BioRender.com

Since LCLA1^{S} contained a putative mPTS at its N-terminus, and presence of this sequence was sufficient to target a fluorescent reporter to the ER, we reasoned that the LCLA1 isoforms could be membrane bound. Therefore, a focus was placed on detection of potential transmembrane α-helices [245]. Three different prediction programs were used to identify possible transmembrane (Tm) helices in both LCLA1 isoforms. Predicted Tm helices were then manually assigned based on the following criteria: 1) agreement for Tm prediction by $\geq$2 Tm prediction algorithms, 2) region of overlapping Tm prediction was also predicted α-helix, and 3) length of Tm domain ranged from 17 to 25 residues in length [246]. Presence of N-terminal Trp and/or C-terminal Gly-cap added further support to likelihood of Tm domain [246]. Through this process, 4 Tm helices were identified with high confidence (Figure 56A).

4.4.10 Loss of LCLA1 also affects mitochondrial morphology

Since *lcla1* was derived from *amc12(#6E9)* originally isolated in a screen for mitochondrial CI assembly deficiency, we asked whether CI was affected in the single mutant as in the original double. To address this question, immunoblotting with antibodies against subunits of the CI soluble arm were performed with total protein extracts from *lcla1* and control samples (Figure 42, top panel). Interestingly, abundance of the 78 kDa subunit was decreased in *lcla1* to similar levels as the *amc9* mutant, a null mutant for the nuclear *NUO5* gene encoding the 78 kDa subunit which was introgressed in the same genetic background [72].

Due to the pleiotropic growth deficiency of *lcla1* and the absence of evidence for mitochondrial localization of $LCLA1^{S}$ or $LCLA1^{L}$, it was reasonable to suspect that mitochondrial deficiency may not be limited to CI. To investigate this, we took advantage of the robust mitochondrial fluorescent marker MDH4-CrVENUS-3xFLAG obtained from the *Chlamydomonas* Resource Center [201]. Notably, expression of MDH4-CrVENUS-3xFLAG (and Cre16.g683600-CrVENUS-3xFLAG as cytosolic control) were robust over >2.5 years of monthly propagation since receipt of each strain. Each strain was crossed into *lcla1* and 6 recombinant haploid haploid progeny of each cross observed by confocal fluorescence microscopy. Intriguingly, mitochondrial morphology was drastically altered in the *lcla1* mutant relative to the parental strain expressing MDH4-CrVENUS-3xFLAG (Figure 57-58).

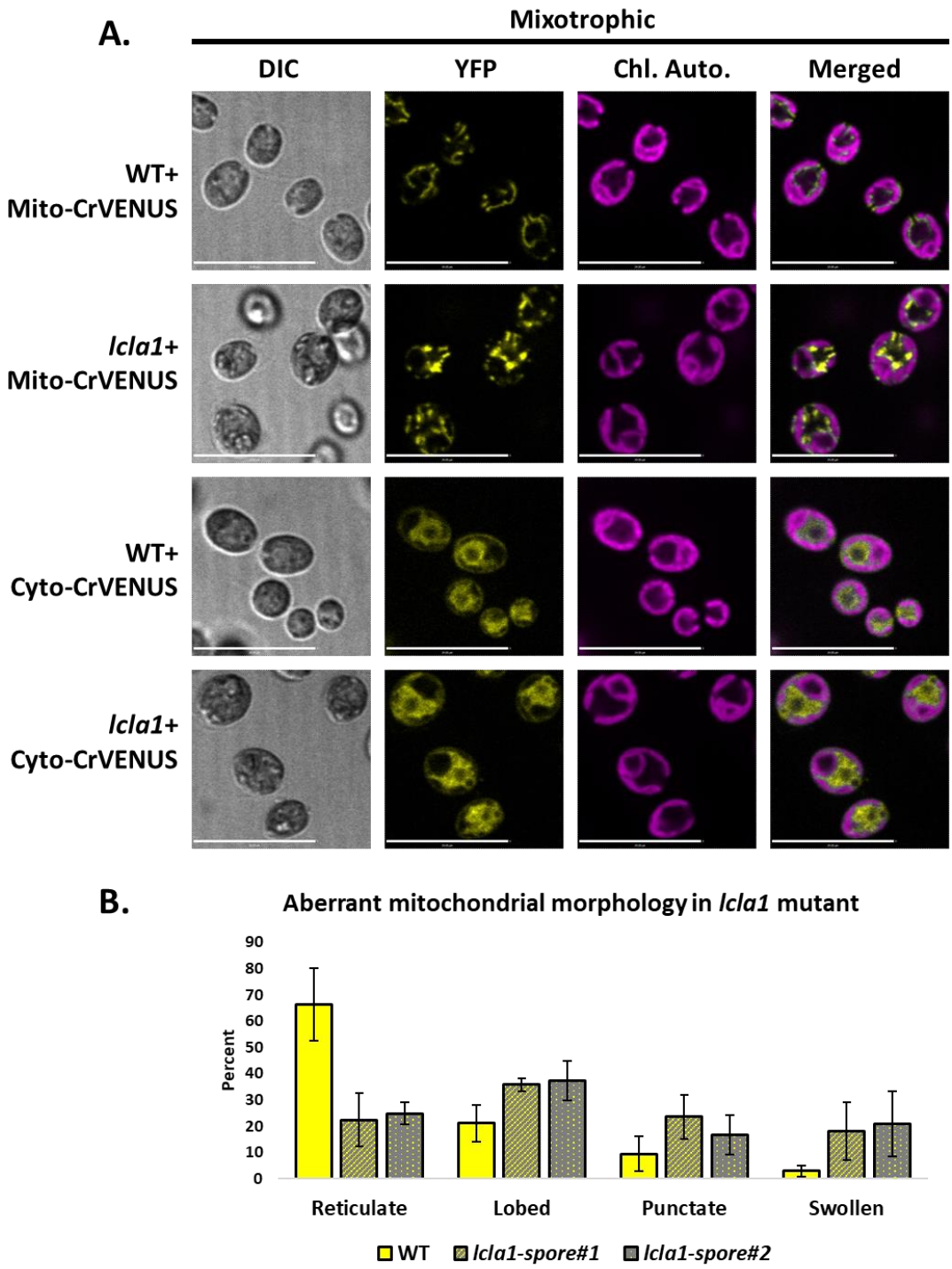


Figure 57 Mitochondrial morphology is affected in the *lcla1* mutant.

Mito-CrVENUS (MDH4-CrVENUS) and Cyto-CrVENUS (Cre16.g683600-CrVENUS) were crossed with *lcla1*, [arg+] (*lcla1#6*) and six recombinant spores of each cross

analyzed. A) Representative confocal microscopy images of mid-exponential phase cultures in mixotrophic conditions (~40 µmol photons $m^{-2}s^{-1}$, 600x total magnification, 150x150 pixels, scale bar is 24 µm). YFP and chlorophyll were excited at 514nm and 640 nm respectively. B) Scoring of mitochondrial morphological defects of two independent spores resulting from cross of *lcla1* x Mito-CrVENUS (4-5 replicates, >250 total cells counted for each sample).

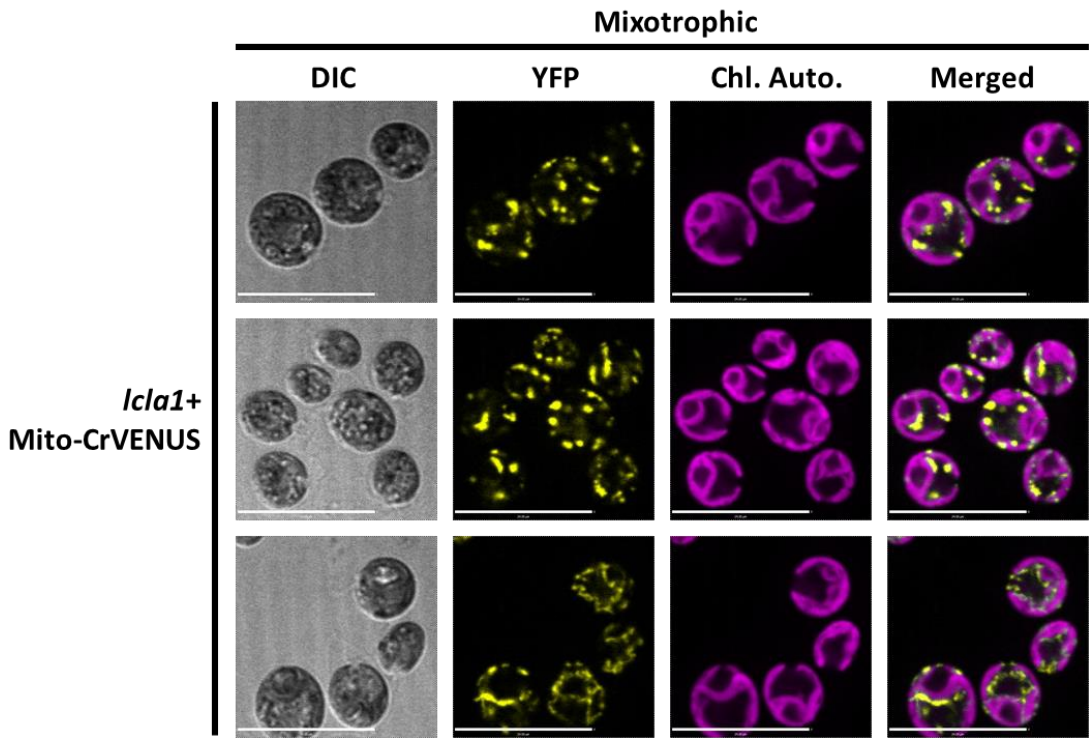


Figure 58 Mitochondrial morphology is affected in the *lcla1* mutant.

Mito-CrVENUS (MDH4-CrVENUS-3xFLAG) was crossed into *lcla1*, [arg+] (*lcla1#6*). Additional representative images showing range of phenotype for lcla1 expressing Mito-CrVENUS in mid-exponential phase, mixotrophic cultures (~40 µmol photons $m^{-2}s^{-1}$, 600x total magnification, 150x150 pixels, scale bar is 24 µm). YFP and chlorophyll were excited at 514nm and 640 nm respectively.

4.5 Discussion

In this chapter, we present a detailed analysis of an insertional mutant (*lcla1*) which was isolated from a genetic screen for CI-deficient mutants (see chapter 2) and displays the characteristic respiratory growth defect (SID) of *amc* strains impaired for CI assembly/activity. In addition to the SID trait but unlike other *amc* mutants analyzed so far, the *lcla1* mutant also exhibits a pleiotropic phenotype with notable growth defect in mixotrophic and phototrophic conditions. The insertional mutation was mapped to the *Cre07.g329861* gene that encodes a large low complexity protein (~325 kDa), which could be immunodetected. While several aspects of the pleiotropic phenotype still remain unexplained, several of the experimental evidence collected during our study suggests acetate assimilation is compromised in the *lcla1* mutant presumably because of a disturbance in peroxisome function, an organelle required for central carbon metabolism [15].

4.5.1 LCLA1 is required for acetate assimilation.

The first indication that the *lcla1* mutant was affected for acetate assimilation was provided by comparison to the *icl* mutant, which is deficient for isocitrate lyase (ICL), a key enzyme of the glyoxylate cycle (Figure 43) [15]. The glyoxylate cycle is a metabolic pathway that allows 2-carbon sources (such as acetate) to be converted into succinate, a C4-molecule to satisfy the cellular carbon requirements. In *lcla1*, the defect in mixotrophic

and respiratory conditions was similar to that of the *icl* mutant with the exception that *icl* is unable to survive in the dark, an expected phenotype since ICL is essential for the flow through the glyoxylate cycle and acetate assimilation required in these conditions (Figure 44A and C) [17]. Because *lcla1* can grow slowly on acetate in the dark and acetate supplementation did not improve the SID phenotype, we reasoned that unlike the *icl* mutant, acetate assimilation still occurs but is restricted. A second observation that reinforced the view that acetate assimilation was impaired due to loss of LCLA1 function came from our genetic studies. Unexpectedly, the mixotrophic growth defect in the *lcla1* was considerably more pronounced when combined with the *arg7-8* mutation (Figure 35B). This is not a trait in other mutant limited for respiratory growth as the mixotrophic growth defect was not observed in the *amc9* mutant that, similarly to *lcla1*, exhibits a SID phenotype (Figures 37 and 44A). (Figure 47B). However, in the *icl* mutant the mixotrophic growth defect is also exacerbated in the presence of the *arg7-8* mutation (Figure 44C) (Figure 44B). Interestingly, analysis of the *icl* mutant grown in mixotrophic conditions showed that while the enzymes of the glyoxylate cycle are down-accumulated, there is a shift of the TCA cycle towards amino acid synthesis including an upregulation of the pathway controlling arginine biosynthesis [15, 17].

The glyoxylate cycle and L-arginine biosynthetic pathway are two well characterized biochemical bypasses of the TCA cycle. Multiple metabolic intermediates of the peroxisome housed glyoxylate cycle are also shared with mitochondria localized TCA cycle (Figure 43). Similarly, *ARG7* encoding arginino-succinate lyase that catalyzes the last step in arginine synthesis, produces fumarate in addition to arginine, thereby

connecting the cycle (Figure 43). Genes encoding proteins involved in similar functions are often co-expressed (Figure 58) [247]. Inspection of publicly available transcriptomic datasets revealed that enzymes of the glyoxylate cycle and those of the arginine biosynthetic pathway are produced in a reciprocal fashion in synchronized cells with genes controlling the glyoxylate cycle expressed at night and genes involved in L-arginine biosynthesis occurring during the day [247]. Worth mention is that synchronization of *Chlamydomonas* cultures is performed in minimal media lacking acetate (Crysten Blaby-Haas, personal communication) which is produced in the dark by fermentative pathways in the chloroplast [16, 21]. So, expression of glyoxylate cycle enzymes is likely induced in response to acetate rather than the absence of light. Indeed, no difference in photosynthetic growth is observed for the *icl*, Arg+ mutant; the difference in growth is only revealed when acetate is present in the media (Figure 44A and C) [15, 17]. It is conceivable that upregulation of the arginine biosynthetic pathway is a response to loss of acetate assimilation in mixotrophic conditions. Considering an *icl* [arg-] strain is considerably affected for mixotrophic growth (compared to its [arg+] counterpart), despite the presence of arginine in the medium, the intracellular synthesis of arginine (or fumarate) must be a requirement for the cell to grow mixotrophically in high light conditions.

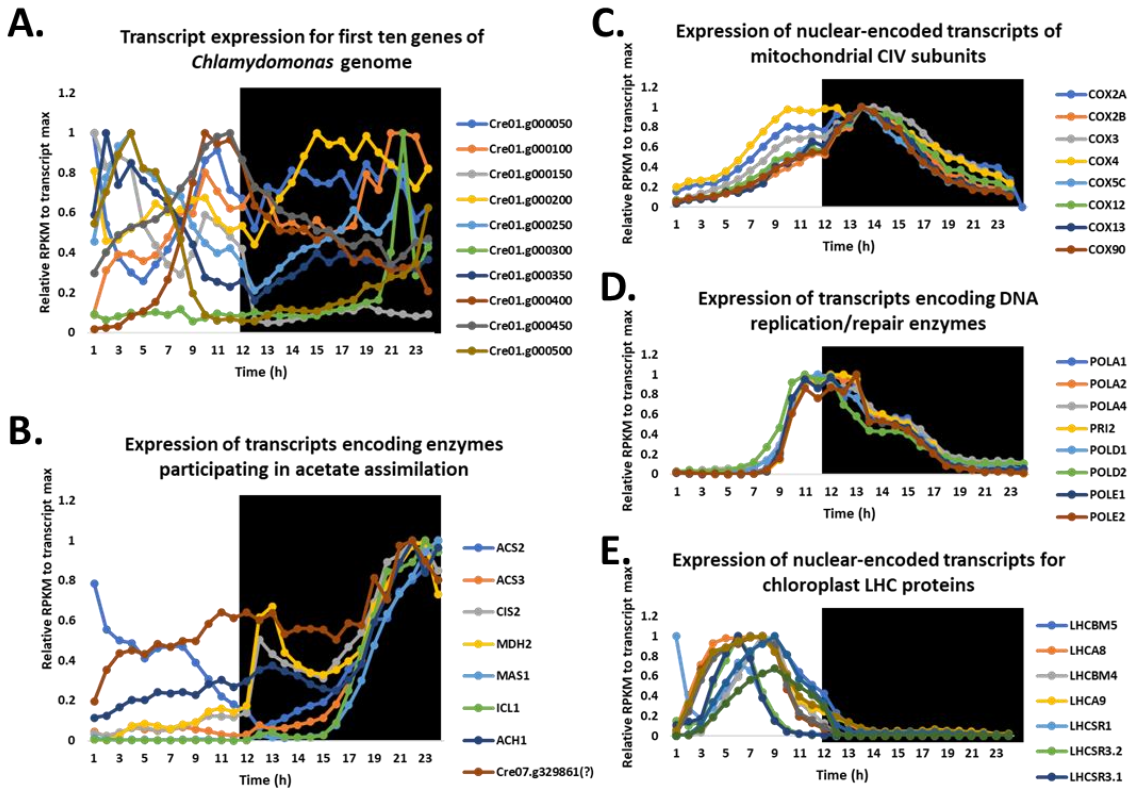


Figure 59 *LCLA1* is co-expressed with enzymes of the glyoxylate cycle.

Re-analysis of publicly available transcriptomic data from synchronized *Chlamydomonas* cultures [247]**.** All transcripts reads were in RPKM (Reads Per Kilobase Million) and normalized to their max value (i.e. each expression pattern is plotted as fraction of its peak value per recommendation of Crysten Blaby-Haas). Black background indicates artificial "night" portion of synchronized cultures from 12/12 hour light/dark cycle. A) Transcript expression profile for first ten genes annotated in the *Chlamydomonas* genome. B) Transcript expression profile for genes encoding enzymes of glyoxylate cycle and participating in acetate assimilation, including possibly LCLA1. C) Transcript expression profile for nuclear genes encoding subunits of mitochondrial CIV. D) Transcript expression profile for enzymes involved in DNA replication and repair. E) Expression pattern of transcripts encoding light harvesting complex proteins.

In *Chlamydomonas,* the L-arginine biosynthetic pathway is entirely plastid localized and expressed during the light phase of diurnally synchronized cultures [247]. The plastid compartmentalization suggests that the L-arginine biosynthetic pathway could help alleviate light stress as a sink for excess electrons in high light mixotrophic conditions; blocking this pathway removes another mechanism for CO_2 fixation (and electron storage) via carbamoyl-phosphate synthase (*Cre08.g358580.t1.1*) which is upstream of ornithine carbamoyltransferase (*Cre12.g489700.t1.2*) in the L-arginine biosynthetic pathway [1, 248-250]. In addition, N-acetyl-γ-glutamyl-phosphate reductase (*Cre03.g146187.t1.1*) regenerates $NADP^+$ [248-250] which can be reused as the terminal electron acceptor of the photosynthetic electron transport chain.

Similar to the *icl* mutant that shifts its metabolism towards fatty acid accumulation in response to loss of acetate assimilation [15, 17], *lcla1* also displays an increase in the abundance of fatty acids in the form of TAGs (Figure 45) and an increased number lipid bodies of larger size, which are specialized storage compartments (Figure 46). All these observations and the finding that *LCLA1* was co-expressed with enzymes of the glyoxylate cycle (Figure 59) led us to postulate that LCLA1 controls acetate assimilation.

4.5.2 LCLA1 is involved in peroxisome biogenesis.

A clue as to the function of LCLA1 in acetate assimilation was provided with the identification of mPTS, a targeting sequence for membrane peroxisomal proteins in the N-terminal extension of $LCLA1^S$, one of the two isoforms (Figures 49). Membrane

peroxisomal proteins (MPPs) are targeted to the peroxisome via trafficking to the ER, following which pre-peroxisomal vesicles containing MPPs form from the ER. This route for peroxisomes biogenesis is referred to as the *de novo* pathway as peroxisomes can also be formed by division of preexisting peroxisomes. A fluorescent reporter containing LCLA1 mPTS localized to the ER during mixotrophic conditions which then redistributed to distinct cytosolic puncta that we hypothesized correspond to peroxisome after 24-hour dark adaptation (Figure 50). This suggests that LCLA1 might be a component of the *de novo* peroxisome biogenesis pathway. A preliminary investigation of the fluorescent reporter expressed in the *lcla1* mutant showed that it fails to redistribute away from the ER (data not shown), a finding compatible with the proposed function of the protein. If the implication of LCLA1 in peroxisome function needs to be further elaborated, several of our observations are compatible with a model where LCLA1 controls peroxisome formation. The impaired acetate assimilation observed in *lcla1* can be explained by loss of the peroxisomal enzymes involved in the glyoxylate cycle, which is also the phenotype displayed by the *icl* mutant. Because the peroxisome is also the site of fatty acid β-oxidation, impaired fatty acid oxidation is expected to be a consequence of a defect in peroxisomal function. Interestingly, we showed that treatment with cerulenin, an inhibitor of plastid fatty acid synthesis alleviates the mixotrophic growth defect in high light of the *lcla1* mutant. This result was taken as an indication the high light sensitivity was due to lipotoxicity due to lipid "overload" (Figure 47B-C). Because this effect of cerulenin is also seen in the *icl* mutant and the abundance of the enzymes involved in fatty acid synthesis is

presumably unchanged, we favor a model where fatty acid are increased because their catabolism is decreased.

That accumulation of fatty acids is due to decreased catabolism because of a defect in peroxisomal function is further supported when investigating whether *lcla1* could utilize its stored carbon for growth. As expected when β-oxidation is impaired, *lcla1* was unable to utilize its own carbon stores for growth compared to wild-type (Figure 48A). In addition, supplementation with oleic acid as the sole carbon source was unable to support growth of the *lcla1* mutant whereas growth of the wild-type strain was boosted after each feeding (Figure 48B-C). Since other forms of stored carbon, most notably starch, are also produced in *Chlamydomonas*, the latter observation that *lcla1* could not utilize oleic acid suggests that the growth phenotype is specific to lipid catabolism, which requires the peroxisome compartment and not due to impaired mobilization of other carbon storage molecules.

Initial homology searches via BLAST of NCBI proteomes only identified potential homologues in the closely related algae *G. pectorale* and *V. carteri* but later searches of the Uniprot database revealed additional proteins of possible interest. One example was *Chlamydomonas Cre08.g384200* encoding a GB1/RHD3-type G domain-containing protein (A0A2K3DIA5, with two guanylate biding domains) which had 29.3% sequence identity to an N-terminal segment of LCLA1. An additional protein with some similarity, also to an N-terminal portion of LCLA1 is *Chlamydomonas* VPS54 (Cre02.g08755, A0A2K3E109). VPS54 has a syntaxin domain which is involved in vesicle fusion [251], so it could be possible that LCLA1 is involved in the sequestration and budding of pre-

peroxisomal vesicles from the ER (containing PEX16) and fusing with PEX3 containing vesicles originating from the mitochondria during peroxisome maturation [252].

Using MEME, a program designed to identify short conserved sequences, several small motifs were found in common between each LCLA1 isoform and the top 20 plant BLAST hits submitted [168, 191] (Figures 56 and 57). Other *Chlamydomonas* proteins containing these motifs were identified using FIMO [191] and submitted for gene ontology enrichment analysis. GO enrichment terms generally fell in one of two categories: transcription factors or proteins involved in post-translational modification by (de)ubiquitination. Transcription factors tend to be large proteins (like LCLA1) and the identified motifs (poly-Q and poly-G) are known to facilitate protein:protein interactions. Intriguingly, protein (de)ubiquitination (or similar) was present in the top 10 GO hits for 5 of the remaining 6 motifs. Import of peroxisomal proteins requires de-ubiquitination to recycle receptor proteins (such as PEX5) for peroxisome resident proteins [240]. Since only 13 of the 34 described peroxins (PEX proteins involved in peroxisome biogenesis/function) are identifiable in *Chlamydomonas*, so room exists for LCLA1 to participate in this process.

While the mechanism of action of LCLA1 in peroxisome biogenesis remains unclear, its inclusion in the family of low complexity proteins is interesting as an emerging role of low complexity proteins is in the self-organization of membrane structures such as protrusions and retractions [253]. It is tempting to speculate that LCLA1 performs a role in peroxisome biogenesis either through membrane envelopment/sequestration from the

ER or by facilitating import of peroxisome resident proteins required for peroxisome function (such as enzymes of the glyoxylate cycle).

4.5.3 Mitochondrial and chloroplast morphology are affected in *lcla1*

Owing to the pleiotropic respiratory and high light growth defects it is unsurprising that both mitochondria and chloroplast morphology are affected (Figure 52). However, it is unclear what the directionality is; whether one is impacting the other or each the result of a shared upstream process such as common metabolite or reductant. Some obvious candidates include NADH/NADPH, ATP, and coenzyme A. Indeed, mitochondrial respiratory deficient mutants (such as those affecting CI) were isolated based on decreased photosynthetic efficiency [25, 99]. Intriguingly, the *icl* mutant was also isolated in this same screen. We presently favor that the observed morphological defects are due to increased phospholipid synthesis (i.e. increased membrane production) as a mechanism to store excess fatty acids and avoid lipotoxicity. We look to confirm these structural defects by electron microscopy and further investigate the chloroplast morphological defect by measurement of *in vivo* PSII electron transport rate and photoprotection capacity.

Chapter 5. Perspectives

Excerpts of this chapter have been previously published:

Subrahmanian, N, **Castonguay, AD**, Fatnes, TA, Hamel, PP. **(2020)** *Chlamydomonas reinhardtii* as a plant model system to study mitochondrial complex I dysfunction. *Plant Direct.* **4**: 1– 16; *https://doi.org/10.1002/pld3.200*

Contributions:
Castonguay, AD: wrote this section and performed all experiments for figures.

5.1 *Chlamydomonas* mitochondrial CI mutants as models for human CI deficiency

An overarching goal of this work was to better understand the contribution of different metabolic pathways contributing to plant growth, with particular focus on those that support growth in respiratory conditions (dark + acetate) which requires mitochondrial complex I (CI). CI, also known as NADH:ubiquinone oxidoreductase is the first and largest enzyme complex of the mitochondrial electron transport chain (ETC) with a molecular weight of approximately 1 MDa [28, 36]. The fully assembled complex is L-shaped with a membrane arm embedded in the inner mitochondrial membrane and soluble arm protruding into the mitochondrial matrix. Together, eukaryotic CI is composed of more than 40 subunits encoded by both the mitochondrial and nuclear genomes, as well as 9 non-protein cofactors (8 iron-sulfur clusters and 1 flavin mononucleotide) [28, 36].

Since loss of CI in mammalian model systems can be lethal, we chose to investigate mitochondrial CI deficiency in the model photosynthetic alga, *Chlamydomonas reinhardtii*. Importantly, *Chlamydomonas* subunit composition is similar to that of human CI [28, 36, 61] and CI mutants are viable but manifest a slow growth in the dark (SID) phenotype when grown in respiratory conditions [8]. In a previously conducted, forward genetic screen of approximately 54,000 transformants, 15 *amc* mutants (for assembly of mitochondrial complex I) were isolated based on the SID phenotype [8, 72, 132]. The CI assembly defect was then confirmed by native gel electrophoresis/in-gel activity staining for mitochondrial CI, and subsequent CI-specific enzymatic activities. Initially, the genetic lesion had only been mapped for one mutant, *amc5*, and found to disrupt the *NUOB10*

nuclear gene encoding the PDSW subunit of CI [8]. In Chapter 2, genetic analyses determined that the SID phenotype was genetically linked to the insertional cassette for 3 additional *amc* mutants: *amc9*, *amc11*, and *amc12*. The corresponding loci of insertions for these mutants were later identified. Whereas the insertional mutations in *amc11* and *amc12* were present in genes for which no prior link to mitochondrial CI had been established, *amc9* represented the second isolated mutant carrying a lesion in CI structural gene [72].

The *amc5* and *amc9* mutants were chosen for further characterization; each mutant was disrupted for a different nuclear gene encoding a CI subunit but importantly corresponded to null alleles (Figures 6C and 13C) [72]. This presented the opportunity to reintroduce mutant versions of each gene to mimic the CI-causing mutations in patients in order to dissect their potential contribution to CI activity or assembly. Introduction of the C79S and/or C91S mutations into the *Chlamydomonas NUOB10* gene (Figure 12) and expression of the mutant versions in *amc5 (nuob10-null)* only partially rescued the SID phenotype (Figure 15A). Similarly, expression of NUOB10 containing either or both substitutions partially restored assembly of holoenzyme with only residual accumulation of the 700 kDa subcomplex observed in the *nuob10-null* recipient strain (Figure 15C) but without any increase in CI specific activity (Figure 15B).

However, a provisional mutation identified in human *NUO5* (K230R) had no impact on CI assembly and activity when reconstructed in *amc9 (nuo5-null)* (Figure 10). This mutation may cause CI disease in humans by disruption of a salt bridge important for protein stability (since arginine and lysine each have basic R-groups, charge conservation is unaffected). In animals, fungi, and bacteria, the adjacent downstream amino acid, an

alanine, is also conserved (Figure 8). In contrast, there are 2 adjacent lysine residues at this position in plant model organisms (*Arabidopsis thaliana*, *Vitis vinifera*, and *Chlamydomonas reinhardtii*, Figure 8); one possibility is that this second lysine can compensate for substitution of the first. Taken together, we concluded that *Chlamydomonas reinhardtii* can serve as a plant model for human CI disorders but further "humanization" of *Chlamydomonas* proteins may be required to fully reconstitute phenotypes observed in human patients. This study, combined with previous work by the Cardol and Remacle laboratories underscores the utility of using *Chlamydomonas* as plant model for human CI disease [22, 72].

5.2 AMC1, a large, low complexity protein is required for expression of mitochondrial-encoded *nd4*

Due to the apparent complexity of mitochondrial CI, the assembly process of the holoenzyme requires proteins not retained in the fully assembled complex that have been collectively termed assembly or biogenesis factors [41, 124, 158]. However, the precise mechanism of assembly and identity of all assembly factors involved is not currently known. To address this knowledge gap, we previously performed a forward genetic screen utilizing *Chlamydomonas reinhardtii* for mutants deficient in CI biogenesis (described in Chapter 2 and 3). Two mutants were isolated and found to carry allelic mutations in *Cre16.g688900* (hereafter *AMC1*, Figure 16A-C). The *amc1-1* and *amc1-2* mutants are significantly affected for respiratory growth due to a specific block in CI assembly, and accumulate a 700 kDa subcomplex, a hallmark of impaired assembly of CI membrane arm

(Figures 17-21). *AMC1* is poorly conserved within related algae, notably *Volvox carteri* and *Gonium pectorale*, and encodes a large protein with numerous regions of biased amino acid composition, belonging to the diverse family of so-called low complexity proteins (Figures 26-27). The N-terminal sequence of AMC1 directed a reporter to yeast mitochondria, a finding that we interpreted as evidence for mitochondrial localization of the protein (Figure 23). Since large proteins of biased amino acid composition (mainly pentatricopeptide and octatricopeptide repeat proteins) have been implicated in chloroplast gene expression, we asked whether AMC1 could be performing a similar function in plant mitochondria. Indeed, in *amc1* mutants, a specific decrease in the mitochondrially encoded *nd4* transcript was observed and this defect was rescued in complemented strains (Figure 24). *Chlamydomonas* mutants in the mitochondrially-encoded *nd4* gene accumulate the same 700 kDa subcomplex identified in *amc1* mutants [22, 24, 63]. Hence if AMC1 controls the expression of *nd4*, loss of AMC1 function is expected to produce the same phenotype (Figure 21A). Importantly, this decrease in the *nd4* transcript was not common to other amc mutants showing an accumulation of a 700 kDa subcomplex such as the *amc5* mutant (Figure 25). The current model is that AMC1 binds directly to the *nd4* transcript to promote its translation. However, attempts to show a decrease in the abundance in the corresponding protein (ND4) with available commercial antibodies in *amc1* mutants were unsuccessful.

Therefore, the most pressing line of experimentation is to test for a direct interaction between AMC1 and its target transcript *nd4* via immunoprecipitation (followed by detection of the transcript) and/or the electrophoresis mobility shift assay with a

recombinant form of the protein. This would require the development of an antibody against the native AMC1 protein, or engineering of a functional AMC1 with a C-terminal epitope tag (such as the 3xFLAG tag) [254].

An interesting observation is that 2 of the 13 *amc* mutants were found to be allelic, harboring different genetic lesions in exon 4 of the *AMC1* gene (Figure 16A-C). We provide evidence in Chapter 3 that AMC1 functions in the expression of the mitochondrial encoded *nd4* transcript, but to date, no other proteins have been implicated in the expression of the 4 remaining mitochondrial encoded CI subunits (*nd5*, *nd6*, *nd2*, and *nd1*). The screen for CI assembly mutants is not saturated and additional *AMC* loci could be revealed.

If the library of CI assembly mutants was to be expanded, several possible modifications should be considered. First, the WT background strains used in this study (4C-, 3A+) do not have as robust growth in the dark compared to other common laboratory strains, notably, CC-4533 and its isogenic mating partner CC-5155 (data not shown). If growth in the dark is already less robust, strains with intermediate phenotypes may be harder to detect and inadvertently missed when scored. Unexpectedly, the latter WT strains have very low complex I activity (measured on crude membrane fractions) compared to 4C- and 3A+; yet, other screens have successfully used CC-125 and 1’ (a laboratory derivative of 137C+) each with CI specific activity similar to CC-4533 and CC-5155 (data not shown).

A second possible modification would be to perform the screening in liquid culture in a 96-well plate format so culture optical density could be monitored to obtain

quantitative differences between mixotrophic and respiratory growth (rather than qualitative assessment of growth on plates). Lastly, carrier DNA was used in the original mutagenesis because it was known to increase transformation efficiency by acting as a decoy for cellular nucleases [90]. However, it was later found to integrate in the *Chlamydomonas* nuclear genome causing random insertions, possibly causal for the CI deficiency observed in the 9 unlinked *amc* mutants [8, 72, 90, 132] (Table 1). Due to this observation, more recent screens have stopped using carrier DNA and instead increased amount of transforming DNA without significant decrease in transformation efficiency [90]. In summary, expansion of this screen, possibly including some or all of the above modifications, would be expected to yield additional CI assembly mutants - not limited to nuclear encoded subunits but also possibly in genes encoding yet to be discovered assembly factors.

5.3 *LCLA1* encodes large protein participating in algal lipid metabolism

In Chapter 4, the *lcla1* mutant was isolated from the original *amc12* mutant (#6E9) obtained in the insertional mutagenesis screen for CI assembly deficient strains [72]. The *amc12* strain manifested a pleiotropic respiratory complex deficiency, with defects in several respiratory enzymes [72], contributing to its SID phenotype (Figure 1). However, a second, independently segregating mutation was identified through genetic analyses contributing to the SID phenotype (Table 1 and unpublished). The first mutation, genetically linked to the insertional cassette marker then became referred to as *lcla1*

whereas the second, unlinked mutation known as "*sid2*" (unpublished). The *sid2* mutation segregated as a single trait and caused a SID phenotype.

Through isolation of *lcla1* away from *sid2*, it became apparent that *lcla1* was affected not only for respiratory growth but also displays a mixotrophic growth defect and a phototrophic growth defect. Since *lcla1* was more affected for phototrophic growth than the original double mutant (*lcla1 sid2*), we reasoned that the *sid2* mutation may actually be a suppressor of the *lcla1* phototrophic growth (and possibly high light sensitivity). This is supported by the observation that *amc12#6E9* (*lcla1 sid2*) grew better photo-autotrophically than the WT strain used for insertional mutagenesis (Figure 60). To directly address this question, we have backcrossed the isolated *sid2* mutant into the same background as *lcla1* and will perform tetrad analysis from zygotes isolated from the cross of *lcla1* x *sid2*. If *sid2* is indeed a suppressor of the *lcla1* mutation, we would predict the mixotrophic and phototrophic defects of *lcla1 sid2* haploid progeny to be suppressed (but not respiratory). If so, it would be of high importance to map the *sid2* mutation to identify the defective gene.

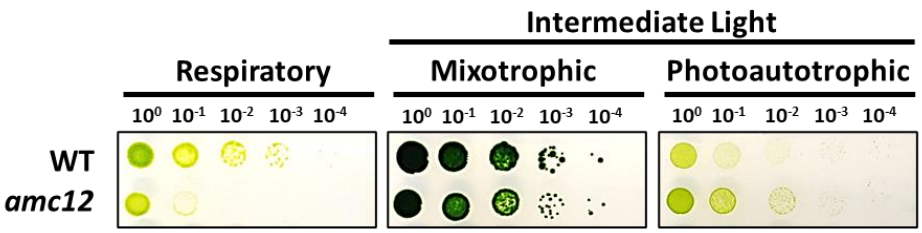


Figure 60 *amc12* original mutant improved for photoautotrophic growth.

Representative images of tenfold dilution series of WT (4C-) or amc12#6E9 (*lcla1 sid2*) in different growth conditions. TARG media was used for both mixotrophic and respiratory growth. Minimal media lacking acetate was used to assess photoautotrophic growth in the same light intensity (~30 µmol photons $m^{-2}s^{-1}$). All pictures taken at 10 days.

To gain a full understanding of LCLA1 function, subcellular localization for each isoform needs to be ascertained. Unfortunately, we have only been able to investigate localization for the LCLA1^{S}-CrVENUS reporter. The N-terminal sequence of LCLA1^{S} contains features resembling the mPTS (peroxisomal targeting sequence) of membrane localized peroxisomal proteins and was sufficient to target a fluorescent reporter to a cellular compartment we identified as the ER (Figures 49-51) [201]. Shifting growth from mixotrophic to respiratory conditions caused the fluorescent reporter to redistribute to a pattern like peroxisomes (Figure 50B). However, co-localization of this fluorescent reporter with known peroxisomal markers (for instance carrying a different fluorophore such as an RFP such as mCherry) remains to be shown. Co-localization of fluorescent markers are difficult in *Chlamydomonas* due to excitation of chlorophyll (and autofluorescence) in multiple channels leading to high background signal [201, 255]. Indeed, high signal of fluorophore to "noise" from chlorophyll autofluorescence [ratio] is required and initial attempts to detect cytoplasmic mCherry were unsuccessful. Similarly, transformation of WT strains with constructs expressing the LCLA1^{L}-CrVENUS isoform was inefficient, yielding only a few transformants (data not shown). Investigation of these transformants by confocal fluorescence microscopy showed aberrant cellular morphology and no chlorophyll autofluorescence, indicating that this chimeric construct produced a toxic protein (data not shown).

Taken together, our experimental evidence is consistent with a defect in acetate metabolism, possibly due to impaired peroxisomal function in the *lcla1* mutant. The *lcla1* mutant phenocopies the *icl* mutant defective in acetate assimilation because the glyoxylate

cycle, which is primarily housed in peroxisomes in *Chlamydomonas* is disrupted (Figures 43-45). This evidence combined with the predicted and observed subcellular localization suggests a role of LCLA1^{L} and LCLA1^{S} in peroxisome biogenesis. Although we were unable to express a reporter to test the cellular localization of the longer isoform, it contains an N-terminal motif similar to the PTS2 of ACS3, a peroxisomal matrix protein. However, a direct connection to peroxisome function still remains to be demonstrated.

To further document the implication of LCLA1 in peroxisome biogenesis, we will undertake several experimental approaches. First, since expression of LCLA1^{L}-CrVENUS was not well tolerated by *Chlamydomonas*, an alternative strategy using yeast as a model system is in development. *S. cerevisiae* mutants deficient in peroxisome function are viable but are sensitive to oleic acid supplemented in the media [200, 256]. Since LCLA1^{L} contains a putative PTS2, and the function of PTS2 is conserved across organisms, we reasoned that a yeast [PTS2-containing] peroxisome resident protein in which the PTS2 has been replaced by that of LCLA1^{L} should still be imported in the peroxisome. To test this, we will use Fox3p (also called Pot1p), an enzyme participating in fatty acid oxidation that is targeted to the peroxisome via its N-terminal PTS2 [200, 256]. Yeast mutants deficient in peroxisome function are viable but are sensitive to oleic acid supplemented in the media [200, 256]. If LCLA1^{L} contains a bonafide PTS2, introduction of this motif in place of the Fox3p PTS2 should rescue oleic acid sensitivity when the chimeric protein is expressed in a *fox3*-null mutant. Alternatively, since the PTS2 sequence is recognized by the peroxisome import proteins Pex7p and Pex5p, a yeast two hybrid assay could be used to on test if LCLA1^{L} PTS2 is interacting with the components of the import machinery

[243]. Similar yeast functional assays could be devised to confirm peroxisomal localization for LCLA1^{S}.

Second, if *de novo* peroxisome biogenesis is impaired in the *lcla1* mutant, a decrease in the number of peroxisomes or peroxisomes with aberrant morphology is an expected phenotype. To our knowledge, there is no peroxisome-specific dyes that can be used in *Chlamydomonas,* so instead a genetic approach is being pursued. Previously, investigation of *Chlamydomonas* glyoxylate cycle enzymes identified N- and C-terminal sequences that were sufficient to target chimeric fluorescent reporters to the peroxisome [15]. A strain expressing one of these fusion proteins, containing the N-terminal 25 amino acids of CIS2 (peroxisomal citrate synthase) including its PTS2 signal (a gift from Dr. Kyle J. Lauersen and Dr. Thomas Baier, formerly of the University of Bielefeld) is available to us (Figure 51, top panel). This strain is currently being crossed with the *lcla1* mutant and recombinant haploid progeny scored for expression of the peroxisome fluorescent marker. We expect that the typical pattern that identifies the peroxisomal compartment with be altered in the *lcla1* mutant (Figure 51 top panel).

Lastly, if peroxisome *de novo* biogenesis is defective in *lcla1*, a decrease in the abundance of peroxisome resident proteins is to be expected. In the *icl* mutant, the glyoxylate cycle enzymes, namely MAS1, MDH1, CIS2, and ACH1 are down-accumulated in response to loss of ICL [15, 17]. Only a few antibodies are presently available that are specific to or cross react with *Chlamydomonas* peroxisome resident proteins. Two of which, α-CAT (catalase) and α-HPR (hydroxypyruvate reductase) are available in our laboratory and will be tested.

Bibliography

1. Hoober, J.K., *The Chlamydomonas Sourcebook. A Comprehensive Guide to Biology and Laboratory Use. Elizabeth H. Harris. Academic Press, San Diego, CA, 1989. xiv, 780 pp., illus. $145.* Science, 1989. **246**(4936): p. 1503-4.
2. Blaby, I.K., et al., *The Chlamydomonas genome project: a decade on.* Trends Plant Sci, 2014. **19**(10): p. 672-80.
3. Merchant, S.S., et al., *The Chlamydomonas genome reveals the evolution of key animal and plant functions.* Science, 2007. **318**(5848): p. 245-50.
4. Kwan, A.L., et al., *Improving gene-finding in Chlamydomonas reinhardtii:GreenGenie2.* BMC Genomics, 2009. **10**: p. 210.
5. baudelet, P.H., et al., *A new insight into cell walls of Chlorophyta.* Algal Research, 2017. **25**: p. 333-371.
6. Salome, P.A. and S.S. Merchant, *A Series of Fortunate Events: Introducing Chlamydomonas as a Reference Organism.* Plant Cell, 2019. **31**(8): p. 1682-1707.
7. Li, X. and M.C. Jonikas, *High-Throughput Genetics Strategies for Identifying New Components of Lipid Metabolism in the Green Alga Chlamydomonas reinhardtii.* Subcell Biochem, 2016. **86**: p. 223-47.
8. Barbieri, M.R., et al., *A forward genetic screen identifies mutants deficient for mitochondrial complex I assembly in Chlamydomonas reinhardtii.* Genetics, 2011. **188**(2): p. 349-58.
9. Dent, R.M., et al., *Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of Chlamydomonas reinhardtii.* Plant Physiol, 2005. **137**(2): p. 545-56.
10. Dent, R.M., et al., *Large-scale insertional mutagenesis of Chlamydomonas supports phylogenomic functional prediction of photosynthetic genes and analysis of classical acetate-requiring mutants.* Plant J, 2015. **82**(2): p. 337-51.
11. Li, X., et al., *An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in Chlamydomonas reinhardtii.* Plant Cell, 2016. **28**(2): p. 367-87.
12. Johnson, X. and J. Alric, *Central carbon metabolism and electron transport in Chlamydomonas reinhardtii: metabolic constraints for carbon partitioning between oil and starch.* Eukaryot Cell, 2013. **12**(6): p. 776-93.
13. Johnson, X. and J. Alric, *Interaction between starch breakdown, acetate assimilation, and photosynthetic cyclic electron flow in Chlamydomonas reinhardtii.* J Biol Chem, 2012. **287**(31): p. 26445-52.

14. Singh, H., et al., *Acetate and bicarbonate assimilation and metabolite formation in Chlamydomonas reinhardtii: a 13C-NMR study.* PLoS One, 2014. **9**(9): p. e106457.
15. Lauersen, K.J., et al., *Peroxisomal microbodies are at the crossroads of acetate assimilation in the green microalga Chlamydomonas reinhardtii.* Algal Research, 2016. **16**(June): p. 266-274.
16. Yang, W., et al., *Algae after dark: mechanisms to cope with anoxic/hypoxic conditions.* Plant J, 2015. **82**(3): p. 481-503.
17. Plancke, C., et al., *Lack of isocitrate lyase in Chlamydomonas leads to changes in carbon metabolism and in the response to oxidative stress under mixotrophic growth.* Plant J, 2014. **77**(3): p. 404-17.
18. Hayashi, Y., et al., *Increase in peroxisome number and the gene expression of putative glyoxysomal enzymes in Chlamydomonas cells supplemented with acetate.* J Plant Res, 2015. **128**(1): p. 177-85.
19. Kong, F., et al., *Chlamydomonas carries out fatty acid beta-oxidation in ancestral peroxisomes using a bona fide acyl-CoA oxidase.* Plant J, 2017. **90**(2): p. 358-371.
20. Durante, L., et al., *Characterization of the GPR1/FUN34/YaaH protein family in the green microalga Chlamydomonas suggests their role as intracellular membrane acetate channels.* Plant Direct, 2019. **3**(6): p. e00148.
21. Yang, W., et al., *Alternative acetate production pathways in Chlamydomonas reinhardtii during dark anoxia and the dominant role of chloroplasts in fermentative acetate production.* Plant Cell, 2014. **26**(11): p. 4499-518.
22. Cardol, P., et al., *In Chlamydomonas, the loss of ND5 subunit prevents the assembly of whole mitochondrial complex I and leads to the formation of a low abundant 700 kDa subcomplex.* Biochim Biophys Acta, 2008. **1777**(4): p. 388-96.
23. Cardol, P., et al., *ND3 and ND4L subunits of mitochondrial complex I, both nucleus encoded in Chlamydomonas reinhardtii, are required for activity and assembly of the enzyme.* Eukaryot Cell, 2006. **5**(9): p. 1460-7.
24. Cardol, P., R.F. Matagne, and C. Remacle, *Impact of mutations affecting ND mitochondria-encoded subunits on the activity and assembly of complex I in Chlamydomonas. Implication for the structural organization of the enzyme.* J Mol Biol, 2002. **319**(5): p. 1211-21.
25. Massoz, S., et al., *Isolation of Chlamydomonas reinhardtii mutants with altered mitochondrial respiration by chlorophyll fluorescence measurement.* J Biotechnol, 2015. **215**: p. 27-34.
26. Massoz, S., et al., *Inactivation of genes coding for mitochondrial Nd7 and Nd9 complex I subunits in Chlamydomonas reinhardtii. Impact of complex I loss on respiration and energetic metabolism.* Mitochondrion, 2014. **19 Pt B**: p. 365-74.
27. Remacle, C., et al., *Mutants of Chlamydomonas reinhardtii deficient in mitochondrial complex I: characterization of two mutations affecting the nd1 coding sequence.* Genetics, 2001. **158**(3): p. 1051-60.
28. Subrahmanian, N., C. Remacle, and P.P. Hamel, *Plant mitochondrial Complex I composition and assembly: A review.* Biochim Biophys Acta, 2016. **1857**(7): p. 1001-14.

29. Salinas, T., et al., *Respiratory-deficient mutants of the unicellular green alga Chlamydomonas: a review.* Biochimie, 2014. **100**: p. 207-18.
30. Green, D.E. and A. Tzagoloff, *The mitochondrial electron transfer chain.* Arch Biochem Biophys, 1966. **116**(1): p. 293-304.
31. Mitchell, P., *Chemiosmotic coupling in oxidative and photosynthetic phosphorylation.* Biol Rev Camb Philos Soc, 1966. **41**(3): p. 445-502.
32. Kerscher, S., et al., *The three families of respiratory NADH dehydrogenases.* Results Probl Cell Differ, 2008. **45**: p. 185-222.
33. Zickermann, V., et al., *Architecture of complex I and its implications for electron transfer and proton pumping.* Biochim Biophys Acta, 2009. **1787**(6): p. 574-83.
34. Hirst, J., *Mitochondrial complex I.* Annu Rev Biochem, 2013. **82**: p. 551-75.
35. Hirst, J., et al., *The nuclear encoded subunits of complex I from bovine heart mitochondria.* Biochim Biophys Acta, 2003. **1604**(3): p. 135-50.
36. Cardol, P., *Mitochondrial NADH:ubiquinone oxidoreductase (complex I) in eukaryotes: a highly conserved subunit composition highlighted by mining of protein databases.* Biochim Biophys Acta, 2011. **1807**(11): p. 1390-7.
37. Berrisford, J.M., R. Baradaran, and L.A. Sazanov, *Structure of bacterial respiratory complex I.* Biochim Biophys Acta, 2016. **1857**(7): p. 892-901.
38. Remacle, C., et al., *Eukaryotic complex I: functional diversity and experimental systems to unravel the assembly process.* Mol Genet Genomics, 2008. **280**(2): p. 93-110.
39. Kmita, K. and V. Zickermann, *Accessory subunits of mitochondrial complex I.* Biochem Soc Trans, 2013. **41**(5): p. 1272-9.
40. Stroud, D.A., et al., *Accessory subunits are integral for assembly and function of human mitochondrial complex I.* Nature, 2016. **538**(7623): p. 123-126.
41. Guerrero-Castillo, S., et al., *The Assembly Pathway of Mitochondrial Respiratory Chain Complex I.* Cell Metab, 2017. **25**(1): p. 128-139.
42. Ghezzi, D. and M. Zeviani, *Human diseases associated with defects in assembly of OXPHOS complexes.* Essays Biochem, 2018. **62**(3): p. 271-286.
43. Rodenburg, R.J., *Mitochondrial complex I-linked disease.* Biochim Biophys Acta, 2016. **1857**(7): p. 938-45.
44. J.A. Morgan-Hughes, P.D., D.N. Landon, J.M. Land, J.B. Clark,, *A mitochondrial myopathy with a deficiency of respiratory chain NADH-CoQ reductase activity.* Journal of the Neurological Sciences, 1979. **43**(1): p. 27-46.
45. Fiedorczuk, K. and L.A. Sazanov, *Mammalian Mitochondrial Complex I Structure and Disease-Causing Mutations.* Trends Cell Biol, 2018. **28**(10): p. 835-867.
46. Friederich, M.W., et al., *Mutations in the accessory subunit NDUFB10 result in isolated complex I deficiency and illustrate the critical role of intermembrane space import for complex I holoenzyme assembly.* Hum Mol Genet, 2017. **26**(4): p. 702-716.
47. Koopman, W.J., et al., *Mitochondrial disorders in children: toward development of small-molecule treatment strategies.* EMBO Mol Med, 2016. **8**(4): p. 311-27.

48. Pagniez-Mammeri, H., et al., *Mitochondrial complex I deficiency of nuclear origin I. Structural genes.* Mol Genet Metab, 2012. **105**(2): p. 163-72.
49. Pagniez-Mammeri, H., et al., *Mitochondrial complex I deficiency of nuclear origin II. Non-structural genes.* Mol Genet Metab, 2012. **105**(2): p. 173-9.
50. Sharma, L.K., J. Lu, and Y. Bai, *Mitochondrial respiratory complex I: structure, function and implication in human diseases.* Curr Med Chem, 2009. **16**(10): p. 1266-77.
51. Distelmaier, F., et al., *Mitochondrial complex I deficiency: from organelle dysfunction to clinical disease.* Brain, 2009. **132**(Pt 4): p. 833-42.
52. Giachin, G., et al., *Dynamics of Human Mitochondrial Complex I Assembly: Implications for Neurodegenerative Diseases.* Frontiers in Molecular Biosciences, 2016. **3**(43).
53. Shoubridge, E.A., *Nuclear genetic defects of oxidative phosphorylation.* Hum Mol Genet, 2001. **10**(20): p. 2277-84.
54. Mitchell, A.L., et al., *Sequence variation in mitochondrial complex I genes: mutation or polymorphism?* J Med Genet, 2006. **43**(2): p. 175-9.
55. Vinothkumar, K.R., J. Zhu, and J. Hirst, *Architecture of mammalian respiratory complex I.* Nature, 2014. **515**(7525): p. 80-84.
56. Lasserre, J.P., et al., *Yeast as a system for modeling mitochondrial disease mechanisms and discovering therapies.* Dis Model Mech, 2015. **8**(6): p. 509-26.
57. Ahlers, P.M., et al., *Application of the obligate aerobic yeast Yarrowia lipolytica as a eucaryotic model to analyse Leigh syndrome mutations in the complex I core subunits PSST and TYKY.* Biochim Biophys Acta, 2000. **1459**(2-3): p. 258-65.
58. Duarte, M., et al., *Neurospora strains harboring mitochondrial disease-associated mutations in iron-sulfur subunits of complex I.* Genetics, 2005. **171**(1): p. 91-9.
59. Kerscher, S., et al., *Application of the yeast Yarrowia lipolytica as a model to analyse human pathogenic mutations in mitochondrial complex I (NADH:ubiquinone oxidoreductase).* Biochim Biophys Acta, 2004. **1659**(2-3): p. 197-205.
60. Maclean, A.E., V.E. Kimonis, and J. Balk, *Pathogenic mutations in NUBPL affect complex I activity and cold tolerance in the yeast model Yarrowia lipolytica.* Hum Mol Genet, 2018. **27**(21): p. 3697-3709.
61. Cardol, P., et al., *Higher plant-like subunit composition of mitochondrial complex I from Chlamydomonas reinhardtii: 31 conserved components among eukaryotes.* Biochim Biophys Acta, 2004. **1658**(3): p. 212-24.
62. Larosa, V., et al., *Reconstruction of a human mitochondrial complex I mutation in the unicellular green alga Chlamydomonas.* Plant J, 2012. **70**(5): p. 759-68.
63. Remacle, C., et al., *High-efficiency biolistic transformation of Chlamydomonas mitochondria can be used to insert mutations in complex I genes.* Proc Natl Acad Sci U S A, 2006. **103**(12): p. 4771-6.
64. Cardol, P., et al., *Photosynthesis and state transitions in mitochondrial mutants of Chlamydomonas reinhardtii affected in respiration.* Plant Physiol, 2003. **133**(4): p. 2010-20.

65. Lecler, R., et al., *Characterization of an internal type-II NADH dehydrogenase from Chlamydomonas reinhardtii mitochondria.* Curr Genet, 2012. **58**(4): p. 205-16.
66. Remacle, C., et al., *Mutations inactivating mitochondrial genes in Chlamydomonas reinhardtii.* Biochem Soc Trans, 2001. **29**(Pt 4): p. 442-6.
67. Banroques, J., A. Delahodde, and C. Jacq, *A mitochondrial RNA maturase gene transferred to the yeast nucleus can control mitochondrial mRNA splicing.* Cell, 1986. **46**(6): p. 837-44.
68. Dujardin, G., et al., *Long range control circuits within mitochondria and between nucleus and mitochondria. I. Methodology and phenomenology of suppressors.* Mol Gen Genet, 1980. **179**(3): p. 469-82.
69. Silhavy, T.J., Berman, M. L., & Enquist, L. W., *Experiments with gene fusions.* Cold Spring Harbor Laboratory Press, 1984.
70. Shimogawara, K., et al., *High-efficiency transformation of Chlamydomonas reinhardtii by electroporation.* Genetics, 1998. **148**(4): p. 1821-8.
71. Berthold, P., R. Schmitt, and W. Mages, *An engineered Streptomyces hygroscopicus aph 7" gene mediates dominant resistance against hygromycin B in Chlamydomonas reinhardtii.* Protist, 2002. **153**(4): p. 401-12.
72. Subrahmanian, N., et al., *Chlamydomonas reinhardtii as a plant model system to study mitochondrial complex I dysfunction.* Plant Direct, 2020. **4**(2): p. e00200.
73. Kropat, J., et al., *A revised mineral nutrient supplement increases biomass and growth rate in Chlamydomonas reinhardtii.* Plant J, 2011. **66**(5): p. 770-80.
74. Werner, R. and D. Mergenhagen, *Mating Type Determination of Chlamydomonas reinhardtii by PCR.* Plant Molecular Biology Reporter, 1998. **16**(4): p. 295-299.
75. Liu, Y.G., et al., *Efficient isolation and mapping of Arabidopsis thaliana T-DNA insert junctions by thermal asymmetric interlaced PCR.* Plant J, 1995. **8**(3): p. 457-63.
76. Purton, S. and J.D. Rochaix, *Complementation of a Chlamydomonas reinhardtii mutant using a genomic cosmid library.* Plant Mol Biol, 1994. **24**(3): p. 533-7.
77. Schagger, H. and G. von Jagow, *Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form.* Anal Biochem, 1991. **199**(2): p. 223-31.
78. Sambrook, J., E.F. Fritsch, and T. Maniatis, *Molecular cloning: a laboratory manual*. 1989, Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. xxxviii + 1546 pp.
79. Dreyfuss, B.W., et al., *Functional analysis of a divergent system II protein, Ccs1, involved in c-type cytochrome biogenesis.* J Biol Chem, 2003. **278**(4): p. 2604-13.
80. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nat Methods, 2012. **9**(7): p. 671-5.
81. Remacle, C., et al., *Impact of a mutation in the mitochondrial LSU rRNA gene from Chlamydomonas reinhardtii on the activity and the assembly of respiratory-chain complexes.* Curr Genet, 2004. **45**(5): p. 323-30.

82. Newman, S.M., et al., *Transformation of chloroplast ribosomal RNA genes in Chlamydomonas: molecular and genetic characterization of integration events.* Genetics, 1990. **126**(4): p. 875-88.
83. de Montaigu, A., et al., *A soluble guanylate cyclase mediates negative signaling by ammonium on expression of nitrate reductase in Chlamydomonas.* Plant Cell, 2010. **22**(5): p. 1532-48.
84. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.
85. Noor-Mohammadi, S., A. Pourmir, and T.W. Johannes, *Method for assembling and expressing multiple genes in the nucleus of microalgae.* Biotechnol Lett, 2014. **36**(3): p. 561-6.
86. Chen, D.C., B.C. Yang, and T.T. Kuo, *One-step transformation of yeast in stationary phase.* Curr Genet, 1992. **21**(1): p. 83-4.
87. Saint-Georges, Y., et al., *A pathogenic cytochrome b mutation reveals new interactions between subunits of the mitochondrial bc1 complex.* J Biol Chem, 2002. **277**(51): p. 49397-402.
88. Depege, N., S. Bellafiore, and J.D. Rochaix, *Role of chloroplast protein kinase Stt7 in LHCII phosphorylation and state transition in Chlamydomonas.* Science, 2003. **299**(5612): p. 1572-5.
89. Saada, A., et al., *C6ORF66 is an assembly factor of mitochondrial complex I.* Am J Hum Genet, 2008. **82**(1): p. 32-8.
90. Zhang, R., et al., *High-Throughput Genotyping of Green Algal Mutants Reveals Random Distribution of Mutagenic Insertion Sites and Endonucleolytic Cleavage of Transforming DNA.* Plant Cell, 2014. **26**(4): p. 1398-1409.
91. Birrell, J.A., et al., *Investigating the function of [2Fe-2S] cluster N1a, the off-pathway cluster in complex I, by manipulating its reduction potential.* Biochem J, 2013. **456**(1): p. 139-46.
92. Nishioka, K., et al., *Genetic variation of the mitochondrial complex I subunit NDUFV2 and Parkinson's disease.* Parkinsonism Relat Disord, 2010. **16**(10): p. 686-7.
93. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.* Mol Syst Biol, 2011. **7**: p. 539.
94. Verkhovskaya, M.L., et al., *Real-time electron transfer in respiratory complex I.* Proc Natl Acad Sci U S A, 2008. **105**(10): p. 3763-7.
95. Sazanov, L.A. and P. Hinchliffe, *Structure of the hydrophilic domain of respiratory complex I from Thermus thermophilus.* Science, 2006. **311**(5766): p. 1430-6.
96. Sokalingam, S., et al., *A study on the effect of surface lysine to arginine mutagenesis on protein stability and structure using green fluorescent protein.* PLoS One, 2012. **7**(7): p. e40410.
97. Zhu, J., et al., *Structure of subcomplex Ibeta of mammalian respiratory complex I leads to new supernumerary subunit assignments.* Proc Natl Acad Sci U S A, 2015. **112**(39): p. 12087-92.

98. Hall, T.A., *BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.* Nucleic Aciids Symposium Series, 1999. **41**: p. 95-98.
99. Massoz, S., et al., *In vivo chlorophyll fluorescence screening allows the isolation of a Chlamydomonas mutant defective for NDUFAF3, an assembly factor involved in mitochondrial complex I assembly.* Plant J, 2017. **92**(4): p. 584-595.
100. Pulkes, T., et al., *Classical mitochondrial phenotypes without mtDNA mutations: the possible role of nuclear genes.* Neurology, 2003. **61**(8): p. 1144-7.
101. Benit, P., et al., *Mutant NDUFV2 subunit of mitochondrial complex I causes early onset hypertrophic cardiomyopathy and encephalopathy.* Hum Mutat, 2003. **21**(6): p. 582-6.
102. Cameron, J.M., et al., *Exome sequencing identifies complex I NDUFV2 mutations as a novel cause of Leigh syndrome.* Eur J Paediatr Neurol, 2015. **19**(5): p. 525-32.
103. Zhang, J., et al., *Association study on the mitochondrial gene NDUFV2 and bipolar disorder in the Chinese Han population.* J Neural Transm (Vienna), 2009. **116**(3): p. 357-61.
104. Vogel, R.O., J.A. Smeitink, and L.G. Nijtmans, *Human mitochondrial complex I assembly: a dynamic and versatile process.* Biochim Biophys Acta, 2007. **1767**(10): p. 1215-27.
105. Petrungaro, C., et al., *The Ca(2+)-Dependent Release of the Mia40-Induced MICU1-MICU2 Dimer from MCU Regulates Mitochondrial Ca(2+) Uptake.* Cell Metab, 2015. **22**(4): p. 721-33.
106. Gabriel, K., et al., *Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway.* J Mol Biol, 2007. **365**(3): p. 612-20.
107. Herrmann, J.M. and J. Riemer, *Three approaches to one problem: protein folding in the periplasm, the endoplasmic reticulum, and the intermembrane space.* Antioxid Redox Signal, 2014. **21**(3): p. 438-56.
108. Szklarczyk, R., et al., *NDUFB7 and NDUFA8 are located at the intermembrane surface of complex I.* FEBS Lett, 2011. **585**(5): p. 737-43.
109. Agip, A.A., et al., *Cryo-EM structures of complex I from mouse heart mitochondria in two biochemically defined states.* Nat Struct Mol Biol, 2018. **25**(7): p. 548-556.
110. Letts, J.A., et al., *Structures of Respiratory Supercomplex I+III2 Reveal Functional and Conformational Crosstalk.* Mol Cell, 2019. **75**(6): p. 1131-1146 e6.
111. Parey, K., et al., *Cryo-EM structure of respiratory complex I at work.* Elife, 2018. **7**.
112. Zhu, J., K.R. Vinothkumar, and J. Hirst, *Structure of mammalian respiratory complex I.* Nature, 2016. **536**(7616): p. 354-358.
113. Weckbecker, D., et al., *Atp23 biogenesis reveals a chaperone-like folding activity of Mia40 in the IMS of mitochondria.* EMBO J, 2012. **31**(22): p. 4348-58.
114. Stan, T., et al., *Mitochondrial protein import: recognition of internal import signals of BCS1 by the TOM complex.* Mol Cell Biol, 2003. **23**(7): p. 2239-50.

115. Letts, J.A. and L.A. Sazanov, *Gaining mass: the structure of respiratory complex I-from bacterial towards mitochondrial versions.* Curr Opin Struct Biol, 2015. **33**: p. 135-45.
116. Hinkle, P.C., *P/O ratios of mitochondrial oxidative phosphorylation.* Biochim Biophys Acta, 2005. **1706**(1-2): p. 1-11.
117. Watt, I.N., et al., *Bioenergetic cost of making an adenosine triphosphate molecule in animal mitochondria.* Proc Natl Acad Sci U S A, 2010. **107**(39): p. 16823-7.
118. Dudkina, N.V., et al., *Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III.* Proc Natl Acad Sci U S A, 2005. **102**(9): p. 3225-9.
119. Baradaran, R., et al., *Crystal structure of the entire respiratory complex I.* Nature, 2013. **494**(7438): p. 443-8.
120. Zickermann, V., et al., *Structural biology. Mechanistic insight from the crystal structure of mitochondrial complex I.* Science, 2015. **347**(6217): p. 44-9.
121. Peters, K., et al., *A structural investigation of complex I and I+III2 supercomplex from Zea mays at 11-13 A resolution: assignment of the carbonic anhydrase domain and evidence for structural heterogeneity within complex I.* Biochim Biophys Acta, 2008. **1777**(1): p. 84-93.
122. Fassone, E., et al., *FOXRED1, encoding an FAD-dependent oxidoreductase complex-I-specific molecular chaperone, is mutated in infantile-onset mitochondrial encephalopathy.* Hum Mol Genet, 2015. **24**(14): p. 4183.
123. Sanchez-Caballero, L., et al., *Mutations in Complex I Assembly Factor TMEM126B Result in Muscle Weakness and Isolated Complex I Deficiency.* Am J Hum Genet, 2016. **99**(1): p. 208-16.
124. Formosa, L.E., et al., *Building a complex complex: Assembly of mitochondrial respiratory chain complex I.* Semin Cell Dev Biol, 2018. **76**: p. 154-162.
125. Ivanova, A., et al., *A Mitochondrial LYR Protein Is Required for Complex I Assembly.* Plant Physiol, 2019. **181**(4): p. 1632-1650.
126. Sheftel, A.D., et al., *Human ind1, an iron-sulfur cluster assembly factor for respiratory complex I.* Mol Cell Biol, 2009. **29**(22): p. 6059-73.
127. Wydro, M.M., et al., *The evolutionarily conserved iron-sulfur protein INDH is required for complex I assembly and mitochondrial translation in Arabidopsis [corrected].* Plant Cell, 2013. **25**(10): p. 4014-27.
128. Schertl, P., et al., *L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in Arabidopsis thaliana.* J Biol Chem, 2012. **287**(18): p. 14412-9.
129. Schimmeyer, J., R. Bock, and E.H. Meyer, *L-Galactono-1,4-lactone dehydrogenase is an assembly factor of the membrane arm of mitochondrial complex I in Arabidopsis.* Plant Mol Biol, 2016. **90**(1-2): p. 117-26.
130. Braun, H.P., et al., *The life of plant mitochondrial complex I.* Mitochondrion, 2014. **19 Pt B**: p. 295-313.
131. Dorthu, M.P., et al., *Biochemical, genetic and molecular characterization of new respiratory-deficient mutants in Chlamydomonas reinhardtii.* Plant Mol Biol, 1992. **18**(4): p. 759-72.

132. Subrahmanian, N., et al., *Assembly of Mitochondrial Complex I Requires the Low-Complexity Protein AMC1 in Chlamydomonas reinhardtii.* Genetics, 2020.
133. Clarke, C.F., W. Williams, and J.H. Teruya, *Ubiquinone biosynthesis in Saccharomyces cerevisiae. Isolation and sequence of COQ3, the 3,4-dihydroxy-5-hexaprenylbenzoate methyltransferase gene.* J Biol Chem, 1991. **266**(25): p. 16636-44.
134. Purton, S. and J.-D. Rochaix, *Characterisation of the ARG7 gene of Chlamydomonas reinhardtii and its application to nuclear transformation.* European Journal of Phycology, 1995. **30**(2): p. 141-148.
135. Lapaille, M., et al., *Atypical subunit composition of the chlorophycean mitochondrial F1FO-ATP synthase and role of Asa7 protein in stability and oligomycin resistance of the enzyme.* Mol Biol Evol, 2010. **27**(7): p. 1630-44.
136. Hsu, A.Y., et al., *Complementation of coq3 mutant yeast by mitochondrial targeting of the Escherichia coli UbiG polypeptide: evidence that UbiG catalyzes both O-methylation steps in ubiquinone biosynthesis.* Biochemistry, 1996. **35**(30): p. 9797-806.
137. Saada, A., et al., *Mutations in NDUFAF3 (C3ORF60), encoding an NDUFAF4 (C6ORF66)-interacting complex I assembly protein, cause fatal neonatal mitochondrial disease.* Am J Hum Genet, 2009. **84**(6): p. 718-27.
138. Asamizu, E., et al., *Generation of expressed sequence tags from low-CO2 and high-CO2 adapted cells of Chlamydomonas reinhardtii.* DNA Res, 2000. **7**(5): p. 305-7.
139. Kent, W.J., *BLAT--the BLAST-like alignment tool.* Genome Res, 2002. **12**(4): p. 656-64.
140. Karolchik, D., et al., *The UCSC Genome Browser Database.* Nucleic Acids Res, 2003. **31**(1): p. 51-4.
141. Gabilly, S.T., et al., *A novel component of the disulfide-reducing pathway required for cytochrome c assembly in plastids.* Genetics, 2011. **187**(3): p. 793-802.
142. von Heijne, G., *Mitochondrial targeting sequences may form amphiphilic helices.* EMBO J, 1986. **5**(6): p. 1335-42.
143. Habib, S.J., W. Neupert, and D. Rapaport, *Analysis and prediction of mitochondrial targeting signals.* Methods Cell Biol, 2007. **80**: p. 761-81.
144. Emanuelsson, O., et al., *Locating proteins in the cell using TargetP, SignalP and related tools.* Nature Protocols, 2007. **2**(4): p. 953-971.
145. Emanuelsson, O., et al., *Predicting subcellular localization of proteins based on their N-terminal amino acid sequence.* J Mol Biol, 2000. **300**(4): p. 1005-16.
146. Tardif, M., et al., *PredAlgo: a new subcellular localization prediction tool dedicated to green algae.* Mol Biol Evol, 2012. **29**(12): p. 3625-39.
147. Fukasawa, Y., et al., *MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites.* Mol Cell Proteomics, 2015. **14**(4): p. 1113-26.
148. Larosa, V. and C. Remacle, *Transformation of the mitochondrial genome.* Int J Dev Biol, 2013. **57**(6-8): p. 659-65.

149. Gray, M.W. and P.H. Boer, *Organization and expression of algal (Chlamydomonas reinhardtii) mitochondrial DNA.* Philos Trans R Soc Lond B Biol Sci, 1988. **319**(1193): p. 135-47.
150. Salinas-Giege, T., et al., *Polycytidylation of mitochondrial mRNAs in Chlamydomonas reinhardtii.* Nucleic Acids Res, 2017. **45**(22): p. 12963-12973.
151. Boer, P.H. and M.W. Gray, *The URF 5 gene of Chlamydomonas reinhardtii mitochondria: DNA sequence and mode of transcription.* EMBO J, 1986. **5**(1): p. 21-8.
152. Boer, P.H. and M.W. Gray, *Short dispersed repeats localized in spacer regions of Chlamydomonas reinhardtii mitochondrial DNA.* Curr Genet, 1991. **19**(4): p. 309-12.
153. Rackham, O., T.R. Mercer, and A. Filipovska, *The human mitochondrial transcriptome and the RNA-binding proteins that regulate its expression.* Wiley Interdiscip Rev RNA, 2012. **3**(5): p. 675-95.
154. Herrmann, J.M., M.W. Woellhaf, and N. Bonnefoy, *Control of protein synthesis in yeast mitochondria: the concept of translational activators.* Biochim Biophys Acta, 2013. **1833**(2): p. 286-94.
155. Kehrein, K., N. Bonnefoy, and M. Ott, *Mitochondrial protein synthesis: efficiency and accuracy.* Antioxid Redox Signal, 2013. **19**(16): p. 1928-39.
156. Christian, B.E. and L.L. Spremulli, *Preferential selection of the 5'-terminal start codon on leaderless mRNAs by mammalian mitochondrial ribosomes.* J Biol Chem, 2010. **285**(36): p. 28379-86.
157. Gallaher, S.D., et al., *High-throughput sequencing of the chloroplast and mitochondrion of Chlamydomonas reinhardtii to generate improved de novo assemblies, analyze expression patterns and transcript speciation, and evaluate diversity among laboratory strains and wild isolates.* Plant J, 2018. **93**(3): p. 545-565.
158. Mimaki, M., et al., *Understanding mitochondrial complex I assembly in health and disease.* Biochim Biophys Acta, 2012. **1817**(6): p. 851-62.
159. Videira, A. and M. Duarte, *From NADH to ubiquinone in Neurospora mitochondria.* Biochim Biophys Acta, 2002. **1555**(1-3): p. 187-91.
160. Radermacher, M., et al., *The three-dimensional structure of complex I from Yarrowia lipolytica: a highly dynamic enzyme.* J Struct Biol, 2006. **154**(3): p. 269-79.
161. Boulouis, A., et al., *The nucleus-encoded trans-acting factor MCA1 plays a critical role in the regulation of cytochrome f synthesis in Chlamydomonas chloroplasts.* Plant Cell, 2011. **23**(1): p. 333-49.
162. Barkan, A. and I. Small, *Pentatricopeptide repeat proteins in plants.* Annu Rev Plant Biol, 2014. **65**: p. 415-42.
163. Cline, S.G., I.A. Laughbaum, and P.P. Hamel, *CCS2, an Octatricopeptide-Repeat Protein, Is Required for Plastid Cytochrome c Assembly in the Green Alga Chlamydomonas reinhardtii.* Front Plant Sci, 2017. **8**: p. 1306.

164. Viola, S., et al., *MDA1, a nucleus-encoded factor involved in the stabilization and processing of the atpA transcript in the chloroplast of Chlamydomonas.* Plant J, 2019. **98**(6): p. 1033-1047.
165. Rahire, M., et al., *Identification of an OPR protein involved in the translation initiation of the PsaB subunit of photosystem I.* Plant J, 2012. **72**(4): p. 652-61.
166. Manna, S., *An overview of pentatricopeptide repeat proteins and their applications.* Biochimie, 2015. **113**: p. 93-9.
167. Rado-Trilla, N. and M. Alba, *Dissecting the role of low-complexity regions in the evolution of vertebrate proteins.* BMC Evol Biol, 2012. **12**: p. 155.
168. Bailey, T.L., et al., *MEME SUITE: tools for motif discovery and searching.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W202-8.
169. Wostrikoff, K., et al., *TCA1, a single nuclear-encoded translational activator specific for petA mRNA in Chlamydomonas reinhardtii chloroplast.* Genetics, 2001. **159**(1): p. 119-32.
170. Auchincloss, A.H., et al., *Characterization of Tbc2, a nucleus-encoded factor specifically required for translation of the chloroplast psbC mRNA in Chlamydomonas reinhardtii.* J Cell Biol, 2002. **157**(6): p. 953-62.
171. Marx, C., C. Wunsch, and U. Kuck, *The Octatricopeptide Repeat Protein Raa8 Is Required for Chloroplast trans Splicing.* Eukaryot Cell, 2015. **14**(10): p. 998-1005.
172. Sen, S., et al., *Role of histidine interruption in mitigating the pathological effects of long polyglutamine stretches in SCA1: A molecular approach.* Protein Sci, 2003. **12**(5): p. 953-62.
173. Schaefer, M.H., E.E. Wanker, and M.A. Andrade-Navarro, *Evolution and function of CAG/polyglutamine repeats in protein-protein interaction networks.* Nucleic Acids Res, 2012. **40**(10): p. 4273-87.
174. Boulouis, A., et al., *Spontaneous dominant mutations in chlamydomonas highlight ongoing evolution by gene diversification.* Plant Cell, 2015. **27**(4): p. 984-1001.
175. Pollastri, G. and A. McLysaght, *Porter: a new, accurate server for protein secondary structure prediction.* Bioinformatics, 2005. **21**(8): p. 1719-20.
176. Buchan, D.W., et al., *Scalable web services for the PSIPRED Protein Analysis Workbench.* Nucleic Acids Res, 2013. **41**(Web Server issue): p. W349-57.
177. Kelley, L.A., et al., *The Phyre2 web portal for protein modeling, prediction and analysis.* Nat Protoc, 2015. **10**(6): p. 845-58.
178. Dosztanyi, Z., et al., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.* Bioinformatics, 2005. **21**(16): p. 3433-4.
179. Dosztanyi, Z., et al., *The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins.* J Mol Biol, 2005. **347**(4): p. 827-39.
180. van der Lee, R., et al., *Classification of intrinsically disordered regions and proteins.* Chem Rev, 2014. **114**(13): p. 6589-631.
181. Choquet, Y. and F.A. Wollman, *Translational regulations as specific traits of chloroplast gene expression.* FEBS Lett, 2002. **529**(1): p. 39-42.

182. Salinas, T., et al., *Co-evolution of mitochondrial tRNA import and codon usage determines translational efficiency in the green alga Chlamydomonas.* PLoS Genet, 2012. **8**(9): p. e1002946.
183. Efremov, R.G., R. Baradaran, and L.A. Sazanov, *The architecture of respiratory complex I.* Nature, 2010. **465**(7297): p. 441-5.
184. Efremov, R.G. and L.A. Sazanov, *Structure of the membrane domain of respiratory complex I.* Nature, 2011. **476**(7361): p. 414-20.
185. Ott, M. and J.M. Herrmann, *Co-translational membrane insertion of mitochondrially encoded proteins.* Biochim Biophys Acta, 2010. **1803**(6): p. 767-75.
186. Hofhaus, G. and G. Attardi, *Lack of assembly of mitochondrial DNA-encoded subunits of respiratory NADH dehydrogenase and loss of enzyme activity in a human cell mutant lacking the mitochondrial ND4 gene product.* EMBO J, 1993. **12**(8): p. 3043-8.
187. Karpova, O.V. and K.J. Newton, *A partially assembled complex I in NAD4-deficient mitochondria of maize.* The Plant Journal, 1999. **17**(5): p. 511-521.
188. Ma, D.P., et al., *Nucleotide sequence of cloned nad4 (urf4) gene from Chlamydomonas reinhardtii mitochondrial DNA.* Gene, 1989. **85**(2): p. 363-70.
189. Pratje, E., et al., *Mitochondrial DNA of Chlamydomonas reinhardtii: the ND4 gene encoding a subunit of NADH dehydrogenase.* Curr Genet, 1989. **16**(1): p. 61-4.
190. Costanzo, M.C., et al., *Highly diverged homologs of Saccharomyces cerevisiae mitochondrial mRNA-specific translational activators have orthologous functions in other budding yeasts.* Genetics, 2000. **154**(3): p. 999-1012.
191. Bailey, T.L., et al., *The MEME Suite.* Nucleic Acids Res, 2015. **43**(W1): p. W39-49.
192. Wichmann, J., K.J. Lauersen, and O. Kruse, *Green algal hydrocarbon metabolism is an exceptional source of sustainable chemicals.* Curr Opin Biotechnol, 2019. **61**: p. 28-37.
193. Guschina, I.A. and J.L. Harwood, *Lipids and lipid metabolism in eukaryotic algae.* Prog Lipid Res, 2006. **45**(2): p. 160-86.
194. Li-Beisson, Y., et al., *The lipid biochemistry of eukaryotic algae.* Prog Lipid Res, 2019. **74**: p. 31-68.
195. Boyle, N.R., et al., *Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in Chlamydomonas.* J Biol Chem, 2012. **287**(19): p. 15811-25.
196. Park, J.J., et al., *The response of Chlamydomonas reinhardtii to nitrogen deprivation: a systems biology analysis.* Plant J, 2015. **81**(4): p. 611-24.
197. Li-Beisson, Y., F. Beisson, and W. Riekhof, *Metabolism of acyl-lipids in Chlamydomonas reinhardtii.* Plant J, 2015. **82**(3): p. 504-522.
198. Kao, Y.T., K.L. Gonzalez, and B. Bartel, *Peroxisome Function, Biogenesis, and Dynamics in Plants.* Plant Physiol, 2018. **176**(1): p. 162-177.
199. Pan, R., et al., *Peroxisomes: versatile organelles with diverse roles in plants.* New Phytol, 2020. **225**(4): p. 1410-1427.

200. Lockshon, D., et al., *The sensitivity of yeast mutants to oleic acid implicates the peroxisome and other processes in membrane function.* Genetics, 2007. **175**(1): p. 77-91.
201. Mackinder, L.C.M., et al., *A Spatial Interactome Reveals the Protein Organization of the Algal CO2-Concentrating Mechanism.* Cell, 2017. **171**(1): p. 133-147 e14.
202. Li, X., et al., *A genome-wide algal mutant library and functional screen identifies genes required for eukaryotic photosynthesis.* Nat Genet, 2019. **51**(4): p. 627-635.
203. Hutner, S.H., et al., *Some Approaches to the Study of the Role of Metals in the Metabolism of Microorganisms.* Proceedings of the American Philosophical Society, 1950. **94**(2): p. 152-170.
204. Dutcher, S.K., *Mating and tetrad analysis in Chlamydomonas reinhardtii.* Methods Cell Biol, 1995. **47**: p. 531-40.
205. Zhang, H., P.L. Herman, and D.P. Weeks, *Gene isolation through genomic complementation using an indexed library of Chlamydomonas reinhardtii DNA.* Plant Mol Biol, 1994. **24**(4): p. 663-72.
206. Onishi, M. and J.R. Pringle, *Robust Transgene Expression from Bicistronic mRNA in the Green Alga Chlamydomonas reinhardtii.* G3 (Bethesda), 2016. **6**(12): p. 4115-4125.
207. Kindle, K.L., *High-Frequency Nuclear Transformation of Chlamydomonas reinhardtii.* Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(3): p. 1228-1232.
208. Li, H.H. and S. Merchant, *Degradation of plastocyanin in copper-deficient Chlamydomonas reinhardtii. Evidence for a protease-susceptible conformation of the apoprotein and regulated proteolysis.* J Biol Chem, 1995. **270**(40): p. 23504-10.
209. Liu, B., et al., *Triacylglycerol profiling of microalgae Chlamydomonas reinhardtii and Nannochloropsis oceanica.* Bioresour Technol, 2013. **146**: p. 310-316.
210. Almagro Armenteros, J.J., et al., *DeepLoc: prediction of protein subcellular localization using deep learning.* Bioinformatics, 2017. **33**(21): p. 3387-3395.
211. Briesemeister, S., et al., *SherLoc2: a high-accuracy hybrid method for predicting subcellular localization of proteins.* J Proteome Res, 2009. **8**(11): p. 5363-6.
212. Horton, P., et al., *WoLF PSORT: protein localization predictor.* Nucleic Acids Research, 2007. **35**(suppl_2): p. W585-W587.
213. Cserzo, M., et al., *Prediction of transmembrane alpha-helices in prokaryotic membrane proteins: the dense alignment surface method.* Protein Eng, 1997. **10**(6): p. 673-6.
214. Bernhofer, M., et al., *TMSEG: Novel prediction of transmembrane helices.* Proteins, 2016. **84**(11): p. 1706-1716.
215. Kouza, M., et al., *The GOR Method of Protein Secondary Structure Prediction and Its Application as a Protein Aggregation Prediction Tool.* Methods Mol Biol, 2017. **1484**: p. 7-24.

216. Madeira, F., et al., *The EMBL-EBI search and sequence analysis tools APIs in 2019.* Nucleic Acids Res, 2019. **47**(W1): p. W636-W641.
217. Letunic, I. and P. Bork, *20 years of the SMART protein domain annotation resource.* Nucleic Acids Res, 2018. **46**(D1): p. D493-D496.
218. Letunic, I., S. Khedkar, and P. Bork, *SMART: recent updates, new developments and status in 2020.* Nucleic Acids Res, 2021. **49**(D1): p. D458-D460.
219. Sager, R. and S. Granick, *Nutritional studies with Chlamydomonas reinhardi.* Ann N Y Acad Sci, 1953. **56**(5): p. 831-8.
220. Sussenbach, J.S. and P.J. Strijkert, *Arginine metabolism in Chlamydomonas reinhardi. On the regulation of the arginine biosynthesis.* Eur J Biochem, 1969. **8**(3): p. 403-7.
221. Kirk, D.L. and M.M. Kirk, *Carrier-mediated Uptake of Arginine and Urea by Chlamydomonas reinhardtii.* Plant Physiol, 1978. **61**(4): p. 556-60.
222. Munoz-Blanco, J., J. Hidalgo-Martinez, and J. Cardenas, *Extracellular deamination of L-amino acids by Chlamydomonas reinhardtii cells.* Planta, 1990. **182**(2): p. 194-8.
223. Zuo, Z., et al., *Study of amino acids as nitrogen source in Chlamydomonas reinhardtii.* Phycological Research, 2012. **60**(3): p. 161-168.
224. Calatrava, V., et al., *Nitrogen scavenging from amino acids and peptides in the model alga Chlamydomonas reinhardtii. The role of extracellular l-amino oxidase.* Algal Research, 2019. **38**: p. 101395.
225. Lin, C.H., et al., *Arrestin-related ubiquitin-ligase adaptors regulate endocytosis and protein turnover at the cell surface.* Cell, 2008. **135**(4): p. 714-25.
226. Whelan, W.L., E. Gocke, and T.R. Manney, *The CAN1 locus of Saccharomyces cerevisiae: fine-structure analysis and forward mutation rates.* Genetics, 1979. **91**(1): p. 35-51.
227. Srb, A.M., *[Antibiotic effect of canavanine on yeasts].* C R Hebd Seances Acad Sci, 1954. **239**(5): p. 447-8.
228. McMahon, D. and P. Langstroth, *The effects of canavanine and of arginine starvation on macromolecular synthesis in Chlamydomonas reinhardi.* J Gen Microbiol, 1972. **73**(2): p. 239-50.
229. Schmollinger, S., et al., *Dissecting the heat stress response in Chlamydomonas by pharmaceutical and RNAi approaches reveals conserved and novel aspects.* Mol Plant, 2013. **6**(6): p. 1795-813.
230. Spreitzer, R.J. and L. Mets, *Photosynthesis-deficient Mutants of Chlamydomonas reinhardii with Associated Light-sensitive Phenotypes.* Plant Physiol, 1981. **67**(3): p. 565-9.
231. Barth, J., et al., *The interplay of light and oxygen in the reactive oxygen stress response of Chlamydomonas reinhardtii dissected by quantitative mass spectrometry.* Mol Cell Proteomics, 2014. **13**(4): p. 969-89.
232. Choquet, Y. and O. Vallon, *Synthesis, assembly and degradation of thylakoid membrane proteins.* Biochimie, 2000. **82**(6-7): p. 615-34.

233. Kuras, R. and F.A. Wollman, *The assembly of cytochrome b6/f complexes: an approach using genetic transformation of the green alga Chlamydomonas reinhardtii.* EMBO J, 1994. **13**(5): p. 1019-27.
234. Monde, R.A., et al., *Post-transcriptional defects in tobacco chloroplast mutants lacking the cytochrome b6/f complex.* Plant J, 2000. **21**(1): p. 61-72.
235. Graham, I.A., *Seed storage oil mobilization.* Annu Rev Plant Biol, 2008. **59**: p. 115-42.
236. Goold, H., et al., *Microalgal lipid droplets: composition, diversity, biogenesis and functions.* Plant Cell Rep, 2015. **34**(4): p. 545-55.
237. Terashima, M., et al., *A fluorescence-activated cell sorting-based strategy for rapid isolation of high-lipid Chlamydomonas mutants.* Plant J, 2015. **81**(1): p. 147-59.
238. Sirevag, R. and R.P. Levine, *Fatty acid synthetase from Chlamydomonas reinhardi.* J Biol Chem, 1972. **247**(8): p. 2586-91.
239. Heredia-Martinez, L.G., et al., *Chloroplast Damage Induced by the Inhibition of Fatty Acid Synthesis Triggers Autophagy in Chlamydomonas.* Plant Physiol, 2018. **178**(3): p. 1112-1129.
240. Farre, J.C., et al., *Peroxisome biogenesis, membrane contact sites, and quality control.* EMBO Rep, 2019. **20**(1).
241. Kamigaki, A., et al., *Identification of peroxisomal targeting signal of pumpkin catalase and the binding analysis with PTS1 receptor.* Plant J, 2003. **33**(1): p. 161-75.
242. Kempinski, B., et al., *The Peroxisomal Targeting Signal 3 (PTS3) of the Budding Yeast Acyl-CoA Oxidase Is a Signal Patch.* Front Cell Dev Biol, 2020. **8**: p. 198.
243. Stehlik, T., et al., *Peroxisomal targeting of a protein phosphatase type 2C via mitochondrial transit.* Nat Commun, 2020. **11**(1): p. 2355.
244. Rottensteiner, H., et al., *Peroxisomal membrane proteins contain common Pex19p-binding sites that are an integral part of their targeting signals.* Mol Biol Cell, 2004. **15**(7): p. 3406-17.
245. Koonin, E.V. and M.Y. Galperin, in *Sequence - Evolution - Function: Computational Approaches in Comparative Genomics*. 2003: Boston.
246. Hildebrand, P.W., R. Preissner, and C. Frommel, *Structural features of transmembrane helices.* FEBS Lett, 2004. **559**(1-3): p. 145-51.
247. Zones, J.M., et al., *High-Resolution Profiling of a Synchronized Diurnal Transcriptome from Chlamydomonas reinhardtii Reveals Continuous Cell and Metabolic Differentiation.* Plant Cell, 2015. **27**(10): p. 2743-69.
248. May, P., et al., *Metabolomics- and proteomics-assisted genome annotation and analysis of the draft metabolic network of Chlamydomonas reinhardtii.* Genetics, 2008. **179**(1): p. 157-66.
249. Cunin, R., et al., *Biosynthesis and metabolism of arginine in bacteria.* Microbiol Rev, 1986. **50**(3): p. 314-52.
250. Hudock, G.S., *The pathway of arginine biosynthesis in Chlamydomonas reinhardi.* Biochem Biophys Res Commun, 1962. **9**: p. 551-5.

251. Teng, F.Y., Y. Wang, and B.L. Tang, *The syntaxins.* Genome Biol, 2001. **2**(11): p. REVIEWS3012.
252. Kim, P., *Peroxisome Biogenesis: A Union between Two Organelles.* Curr Biol, 2017. **27**(7): p. R271-R274.
253. McCusker, D., *Cellular self-organization: generating order from the abyss.* Mol Biol Cell, 2020. **31**(3): p. 143-148.
254. Hellman, L.M. and M.G. Fried, *Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions.* Nat Protoc, 2007. **2**(8): p. 1849-61.
255. Yang, W., et al., *Critical role of Chlamydomonas reinhardtii ferredoxin-5 in maintaining membrane structure and dark metabolism.* Proc Natl Acad Sci U S A, 2015. **112**(48): p. 14978-83.
256. Grunau, S., et al., *Peroxisomal targeting of PTS2 pre-import complexes in the yeast Saccharomyces cerevisiae.* Traffic, 2009. **10**(4): p. 451-60.