REGULATION OF CHLAMYDOMONAS REINHARDTII CELL CYCLE DURING SYNCHRONOUS GROWTH

By

Yang-Tsung Lin

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biochemistry and Molecular Biology—Doctor of Philosophy

2022

ABSTRACT

REGULATION OF CHLAMYDOMONAS REINHARDTII CELL CYCLE DURING SYNCHRONOUS GROWTH

By

Yang-Tsung Lin

The regulation of the cell cycle in microalgae has long been a central topic in third-generation biofuel studies, as it relates to biomass accumulation and lipid production, the two determining factors of economically feasible biofuel production. The discovery of Compromised Hydrolysis of TAG 7 (CHT7) in the green alga *Chlamydomonas reinhardtii* has provided valuable insights into the metabolic status-dependent regulation of the cell cycle. CHT7 belongs to the CHC protein family, members of which include transcriptional regulators programming the initiation of cell division. A similar role of CHT7 in mediating cell cycle progression has been suggested. However, the regulatory mechanism and specific processes regulated by CHT7 have yet to be determined. In this dissertation, I applied bioinformatics approaches to comprehensively study CHT7-mediated gene regulation during the cell cycle using cell cultures synchronously grown in bioreactors. I found that specific pathways such as DNA replication and spindle assembly are affected by the absence of CHT7. Other affected pathways include cell-wall remodeling and previously uncharacterized putative kinase cascades. In addition, I discovered the presence of two potential cis-regulatory elements near the transcription start site of misregulated genes in *cht7*, which are potentially linked to homeodomain transcription factors. Moreover, I explored the potential relationship between the CHT7-mediated pathway and the gene network governed by the retinoblastoma protein complex. Lastly, I characterized the function of a novel CHC protein in cell cycle progression and compared it with the role of CHT7. Together, these studies facilitate a better understanding of the cell cycle regulation in microalgae.

ACKNOWLEDGMENTS

Ever since I was young, I have dreamed of becoming a scientist to discover the beauty and secrets of nature. From children's picture books about nature to experimental protocols in the lab, on the path of pursuing my Ph.D., I had always thought it was the curiosity and fascination for science motivating me all along to achieve this goal. However, looking back at the past few years, I realized that it is the people surrounding me and their unfailing support making my dream come true.

I would like to thank my Ph.D. advisor, Dr. Christoph Benning, for being a wonderful mentor throughout my Ph.D. I appreciate his trust in me and always putting my best interest in mind. He is critical, decisive, but also compassionate. He showed me what a true leader is like. I am also grateful to Dr. Barb Sears who always gives me practical advice on research and writing. Outside of work, she encourages me to experience life in the Midwest. Because of her, Michigan feels like a second home to me. I would like to thank our lab manager Linda Danhof for not only keeping the lab organized and safe but also always bringing a delightful and cheerful atmosphere during festivals.

I would like to express my gratitude to my dissertation committee members, Dr. David Arnosti, Dr. Eva Farre, Dr. Hideki Takahashi, and Dr. Min-Hao Kuo. They have been insightful, constructive, and critical of my research but have also been friendly and understanding throughout my Ph.D. Finally, I am thankful to the Benning lab members from 2016 to 2022 for making the lab a fun, collaborative, and enjoyable working environment.

Last but not least, I want to thank my wife Sophie for always being by my side and supporting me unconditionally. I could not have done this without her.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KEY TO ABBREVIATIONS

AA – Amino acid

APC – Anaphase-promoting complex

ARC6-like – Accumulation and Replication of Chloroplasts 6-like

ASF1 – Anti-Silencing Factor 1

ATM – Ataxia Telangiectasia Mutated

ATR – Ataxia Telangiectasia Mutated and Rad3-related

AUG1 – AUGMIN subunit 1

BUB1 – Budding Uninhibited by Benzimidazoles 1

BRCA2 – Breast Cancer gene 2

CAPD2 – Chromosome-Associated Protein D2

CBLP – Chlamydomonas beta-subunit-like polypeptide

CCaMK – Ca^{2+}/CaM-dependent Protein Kinase

CDC20 – Cell Division Cycle 20

CDH1 – Cdc20 Homologue-1

CDC45 – Cell Division Cycle 45

CDK – Cyclin-dependent kinase

Cdt1 – Cdc10-dependent transcript 1

ChIP – Chromatin Immunoprecipitation

CHC – CXC-Hinge-CXC

CHR – Cell cycle genes homology region

CHR1/2 – CHC-related 1/2

CHT7 – Compromised Hydrolysis of Triacylglycerols 7

CIPK – CBL-interacting Protein Kinase

CNK – Chlamydomonas NIMA-related Kinase

Co-IP – Co-immunoprecipitation

CPK – Calcium-dependent Protein Kinases

CREs – cis-regulatory element

CTIP – CtBP Interacting Protein

Cw^- – Cell-wall minus

Cw^+ – Cell-wall plus

CYC – Cyclin

DAPI – 4′,6-diamidino-2-phenylindole

DEG – Differentially Expressed Gene

DP1 – Dimerization partner 1

DREAM – DP, RB, E2F, Myb-MuvB

E2F – Adenovirus early gene 2 binding factor

FIMO – Find Individual Motif Occurrences

FSC – Foward scatter

GCP4 – γ-tubulin Complex Protein 4

HTV1 – Histone H3 variant 1

HLM8 – Histone-lysine N-methyltransferase 8

HOMER – Hypergeometric Optimization of Motif EnRichment

ePBR – Environmental photobioreactor

GAS28 – Gamete-Specific Glycoprotein 28

GFP – Green Fluorescent Protein

GSK3 – Glycogen Synthase Kinase 3

GSM1 – Gamete-Specific Minus Molecule 1

GSP1 – Gamete-Specific Plus Molecule 1

GSL1 – Gibberellin Stimulated-Like protein 1

GO – Gene ontology

H3 – Histone 3

HA – Hemagglutinin

HS – High salt medium

KIN14A-1 – Kinesin-like Protein 14A-1

KNAT – Knotted-Like From Arabidopsis Thaliana

Log_2FC – Log_2 fold change

MCM8 – Minichromosome Maintenance 8

MAT3 – Mating-type protein 3/ RB homolog

MEME – Multiple Em for Motif Elicitation

MID – Minus Dominance

MIND – Mini-cells D

MIP – Myb-interacting protein

MMP– Matrix Metalloprotease

MND1 – Meiotic Nuclear Divisions 1

Mt^- – Mating-type minus

Mt^+ – Mating-type plus

MuvB – Multi-vulva class B

MYB3R – Myb three repeats protein

N – Nitrogen

N+ – N-replete

ND – N-deprived

NR – N-resupplied

NEK – NIMA-related Kinase

NUF2 – Nuclear Filament-containing protein 2

OD – Optical density

ORC1 – Origin Recognition Complex 1

PETO – Cytochrome b6f complex subunit V

PHC – Pherophorin

PL – Parental line

POLA1/2 – DNA polymerase Alpha 1/2

POLD1 – DNA polymerase Delta 1

qPCR – Quantitative PCR

RB – Retinoblastoma protein

RFA1 – Replication Factor A1

RNA-Seq – RNA-Sequencing

RT-qPCR – Real-time quantitative-PCR

S/M phase – DNA synthesis and mitosis phase

SMC4 – Structural Maintenance of Chromosomes 4

SMC5 – Structural Maintenance of Chromosomes 5

SPC25 – Spindle Pole Component 25

SRR – Scavenger Receptor Cysteine-rich

STM – Shoot Meristemless

STREME – Sensitive, Thorough, Rapid, Enriched Motif Elicitation

SUB – Subtilisin-like protease

TOP2 – Topoisomerase II

TAG – Triacylglycerols

TALE – Three Amino acid Loop Extension

TCX5 – TSO1-like CXC 5

TSS – Transcription Start Site

UTR – Untranslated Region

VLE1 – Vegetative Lytic Enzyme 1

WT – Wild type

ZT – Zeitgeber

CHAPTER 1
The study of the cell cycle in green microalgae Chlamydomonas: A literature review

To grow or not to grow: a critical decision

Throughout the life cycle of microorganisms, it is important for cells to progress through the appropriate stages of the cell cycle. As defined by the cell's metabolic activity and cell division, cells are either actively dividing or in a quiescent state. (Fig. 1.1) The former is usually considered as a metabolically active state where the cell generates and utilizes energy efficiently to enlarge and prepare for cell division. In contrast, a cell in quiescence is less active metabolically, with a goal of reducing energy use and waiting for the correct time to re-enter the cell-division cycle (Gray et al., 2004; Valcourt et al., 2012). The ability to transition between the two states could be critical for cells to adapt to a changing environment. For example, when nutrients are limiting, there is a selective advantage for cells to stay in the resting quiescence state rather than undergo cell division. On the other hand, when nutrients are plentiful, it will be more advantageous for the quiescent cells to re-enter the cell-division cycle in order to grow and multiply. Studying the controls of the transition between these two cell states will allow us to understand the survival strategies of microorganisms in response to environmental signals.

Microalgae as a potential feedstock for biofuel production

An important reason for studying the life cycle transitions of unicellular organisms is that during quiescence, cells tend to store energy in the form of organic polymers such as glycogen, starch, and lipids. Among them, oils synthesized by microalgae have long been considered for their high potential as feedstock for third-generation biofuels (Saad et al., 2019; Satyanarayana et al., 2011). Compared to the first- and second-generation biofuels produced from oil crops such as soybean and coconut, algal biofuels have a greater productivity per land area (Fig. 1.2) (Barry et al., 2016; Hallenbeck et al., 2016; Liu and Benning, 2013; Satyanarayana et al., 2011). This advantage is in

part due to the higher lipid content per cell in microalgae. For instance, the green alga *Chlorella vulgaris* can accumulate lipids up to 60% (dry weight) under certain conditions, while soybean seeds contain about 20% (dry weight) of lipids (Dastmalchi, 2021; Guarnieri et al., 2011). However, as mentioned above, oil accumulation in microalgae is often linked to the stress-induced transition of cell cycle states (Boyle et al., 2012; Li-Beisson et al., 2019). To date, one of the main challenges of biofuels from algae is that oil production is usually accompanied by a low biomass accumulation (Khan et al., 2018). It is estimated that the cost of algal oil could drop from $25 per gallon to $2.5 per gallon if the productivity of algal lipids can be increased from 10g $m^{-2}day^{-1}$ at 15% triacylglycerol (TAG) content to 50g $m^{-2}day^{-1}$ at 50% TAG (Pienkos and Darzins, 2009). To achieve this goal, a comprehensive understanding of the regulatory mechanisms involving both lipid biosynthesis and cell cycle progression is necessary.

Cellular responses to nitrogen deprivation and the discovery of CHT7 in Chlamydomonas

Nitrogen (N) deprivation of agal cultures has been an effective approach for laboratory studies of the quiescence state and TAG accumulation. In the model green alga *Chlamydomonas reinhardtii* (hereafter referred to as Chlamydomonas), studies have focused on the global changes in transcripts, proteins, and metabolites in N-deprived (ND) cells. In general, those changes are 1) down-regulation of photosynthesis to prevent the production of cytotoxic ROS, 2) diversion of carbon flux from the TCA/Glyoxylate cycle to carbon storage starch and TAG, and 3) induction of gametogenesis for mating and zygote formation to survive environmental stress (Beck and Acker, 1992; Miller et al., 2010; Park et al., 2015; Schmollinger et al., 2014; Vallon and Wollman, 1995). However, although the survival strategies of Chlamydomonas under N deprivation are well-

studied, the specific regulatory controls of the cell cycle during the transition into and out of quiescence remain largely unknown.

In 2014, a Chlamydomonas mutant generated by random linearized plasmid insertion was found to be impaired in TAG degradation after the N-deprived culture was resupplied with N. The mutated gene responsible for this phenotype was identified and named *COMPROMISED HYDROLYSIS OF TRIACYLGLYCEROLS 7* (*CHT7*, Cre11.g481800) (Tsai et al., 2014). Assessment of growth after N resupply indicated that *cht7* cells were delayed in exiting quiescence (Tsai et al., 2014). The majority of CHT7 protein is located in the nucleus and is part of a protein complex. Later transcriptome analyses confirmed that a subset of quiescence-related genes was misregulated in *cht7* under both ND and N-resupplied (NR) conditions (Tsai et al., 2018; Tsai et al., 2014), indicating that CHT7 is required for cells to program the transition between the the normal cell division cycle and quiescence (Fig. 1.1). Subsequent work in the Benning lab has shown that during N deprivation, the *cht7* mutant gradually loses viability compared to WT and the complementation line *CHT7HA::cht7* (Takeuchi et al., 2020). Furthermore, some *cht7* cells under ND and NR conditions were observed to have multiple nuclei and an abnormal cell shape. Additional transcriptome analysis revealed that genes involved in the S/M phase (S: S-phase, M: Mitosis) are de-repressed in *cht7* cells during N-replete and ND growth (Takeuchi et al., 2020). Together, these findings suggested that CHT7 may transcriptionally regulate the transition between cell cycle states, and the delayed degradation of TAG in N-fed *cht7* cells may be due to the failure of ND cells to resume the normal cell division cycle. This work led our group to explore further the potential involvement of CHT7 in cell cycle progression in Chlamydomonas, which will be the focus of the work described in subsequent chapters.

The synchronized Chlamydomonas cell culture as a platform for cell cycle study

One approach for studying the cell cycle is to see how different genes/proteins are regulated during different stages of their cell cycle. The versatile, green alga Chlamydomonas with a genetic mating system and fully sequenced genome has provided a powerful platform to investigate the gene network governing the cell cycle. The experimental conditions for discriminating the stages of the cell cycle are well-established in Chlamydomonas. When grown under autotrophic and diurnal conditions (typically 12h light / 12h dark), wild-type Chlamydomonas cultures become synchronized, and different stages of the cell cycle can therefore be recognized along the timeline throughout the day (Lien and Knutsen, 1979). The use of a photobioreactor (PBR) further improves studies on the algal cell cycle since it enables the continuous growth of cultures under a stable, programmable, and scalable growth condition (Lucker et al., 2014; Strenkert et al., 2019; Zones et al., 2015). As shown in Figure. 1.3, synchronized Chlamydomonas cells have a prolonged G1 phase when they grow rapidly during the day to enlarge in cellular size. After passing the cell size commitment checkpoint, which usually takes place 1-2 hours before the dark, the cell keeps growing until the end of the day and then enters the S/M phase, when it undergoes multiple rounds of cell division. The number of cell divisions is dependent on the cell size, with larger cells producing more progeny cells. Finally, the newly divided cells will enter the resting G0 phase for the remainder of the dark period and start a new round of the cell-division cycle (G1-S/M) the next day (Zones et al., 2015).

Checks and balances: protein complexes governing the cell cycle in Chlamydomonas

Cell division is an irreversible process involving several critical checkpoints to ensure steps like DNA replication, spindle formation, chromosome segregation, and cytokinesis take place in a

timely manner. In addition, the number of cell divisions during the S/M phase is also heavily regulated in Chlaymdononas to ensure that the resulting daughter cells are uniform in size. Below I briefly review the current knowledge of how cell cycle progression is controlled in Chlamydomonas, which is summarized in Figure 1.4.

In eukaryotes, cyclin-dependent kinases (CDKs) have been considered as the "engine" driving the progression of the cell cycle, while protein inhibitors and U3 ubiquitin ligases like the anaphase-promoting complex (APC) act as the "brake" for the CYC-CDK activity (Salazar-Roa and Malumbres, 2017). In animals and fungi, the primary cell cycle driver CDK1 (also known as CDC2 in animals and CDC28 in yeast) is activated by type-A and B cyclins (CYCs) during the late G1 phase. The CYC-CDK1 complex phosphorylates multiple targets, including the transcriptional repressor Whi5 and the Forkhead transcription activators Fkh1 and Fkh2, which programs the cell to enter the S/M phase (Deshpande et al., 2005; Draetta, 1990; Enserink and Kolodner, 2010). In plants, the CDK1 ortholog CDKA is dispensable, while the plant-specific CDKB family has a CDK1-like function in promoting cell division (Cross and Umen, 2015; Nowack et al., 2012). In Chlamydomonas, there is one copy of CDKA and CDKB each, namely CDKA1 and CDKB1, which are likely bound to CYCA1 and CYCB1, respectively (Atkins and Cross, 2018; Tulin and Cross, 2014). The *cdka-1* mutant exhibits a delayed entry into the S/M phase but is able to eventually complete cell division. The *cdkb-1* mutant, on the other hand, arrests division after the first round of DNA replication, and no spindle formation is observed (Tulin and Cross, 2014). Later studies confirmed that CYCA-CDKA1 regulates the timing of cell division by activating the mitotic transcriptome, including genes for DNA replication and spindle assembly, which might be associated with cell size-dependent commitment (Cross, 2020; Tulin and Cross, 2015). The expression of genes encoding CYCB1 and CDKB1 is dependent on CYCA1-CDKA1,

which in turn inhibits the activity of CDKA1 through a negative feedback loop. In comparison with CYCA1-CDKA1, CYCB1-CDKB1 is essential for spindle formation and is potentially required for the activation of the APC (Atkins and Cross, 2018; Tulin and Cross, 2015). The Chlamydomonas mutant of APC, *cdc27-6*, is arrested during the metaphase when mitotic spindles are present and CDKA1 and CDKB1 are continuously active, suggesting a conserved function of APC in the deactivation of mitotic CDKs by promoting the degradation of corresponding CYCs (Atkins and Cross, 2018; Eloy et al., 2015).

As mentioned earlier, the control of cell size is another important aspect of the cell cycle regulation in Chlamydomonas. There are two major mechanisms adopted by Chlamydomonas cells to make sure each complete cell division cycle generates equally sized daughter population: the cell size checkpoint commitment during the late G1 phase and the maintenance of cell division activity during the S/M phase. The first mechanism determines the threshold to enter cell division (the lower limit or commitment), while the latter decides how many rounds of division a cell undergoes during S/M (the upper limit) (Cross, 2020; Umen, 2005).

In Chlamydomonas, the disruption of the *Mating-type Protein 3* (*MAT3*) gene results in cells with smaller size, a phenotype caused by premature passing of the commitment point and more rounds of division during the S/M phase. *MAT3* encodes the Retinoblastoma (RB) protein which is the only RB homolog in Chlamydomonas (Umen and Goodenough, 2001). As the name suggests, the *rb1* mutant in humans was originally linked to tumors. Later studies showed that RB1 is a master transcriptional repressor that forms a stable complex with the E2 promoter binding factor (E2F) family and dimerization partner (DP) to regulate the progression from G1 to S phase (Engel et al., 2015; Friend et al., 1986). The RB homolog in Arabidopsis, RBR1, is also found to interact with E2F and DP. In addition, the RB-E2F-DP complex is associated with proteins such

as Tesmin/TSO1-like CXC domain protein (TCX5), MULTI-COPY SUPPRESSOR OF IRA1 (MSI1), and different MYBs to form a complex similar to a so-called DREAM complex in animals, which alternatively represses and activates transcription programs for cell cycle progression (Desvoyes and Gutierrez, 2020; Kobayashi et al., 2015). The function of the DREAM complex will be discussed below.

The finding of *mat3* suppressors *dp1* and *e2f* as well as the later protein interaction assays confirmed the presence of the RB-DP-E2F complex in Chlamydomonas. Interestingly, the *dp1* and *ef2* mutants exhibited opposite phenotypes to *mat3*, including larger cells, early entry into the S/M phase, and fewer rounds of cell division (Fang et al., 2006; Olson et al., 2010). The activity of RB was later proven to be negatively regulated by D cyclin-dependent CDKG1 whose protein abundance is positively correlated with the mother cell size (Li et al., 2016). Notably, unlike previously mentioned *cdka1* and *cdkb1*, the activation of genes for DNA replication and mitosis is not affected by the absence of MAT3 (Olson et al., 2010). A recent study also shows that CDKA1 plays a dominant role in determining the commitment size (lower limit), while the MAT3 (RB) complex mainly affects the division number (upper limit). In summary, the Chlamydomonas CDKA1 and RB-DP-E2F appear to regulate cell size by distinct pathways during the cell cycle.

The CHC Proteins and the MuvB complex in Cell Cycle Progression

The CHT7 protein has been localized to the nucleus and contains a CHC (CXC-Hinge-CXC) domain composed of two cysteine-rich CXC motifs (Tsai et al., 2014). The best-known CHC protein in animals is LIN54/Mip120 which was initially isolated from a cell lineage *lin-54* mutant of *Caenorhabditis elegans* and later found in *Drosophila* and named Mip120 (Myb-interacting protein 120). LIN54 is an evolutionarily conserved subunit of the Multi-vulval class B (MuvB)

core complex that plays an essential role in cell cycle regulation (Beall et al., 2002; Sadasivam and DeCaprio, 2013; Thomas et al., 2003). In mammalian cells at the G0 stage, the MuvB core associates with an RB-like protein complex to form the earlier-mentioned DREAM (DP, RB, E2F, Myb-MuvB) complex in order to repress the expression of cell cycle genes. During the G2 phase, MuvB alternatively interacts with BMYB transcription factors to promote entry into the M phase (Litovchick et al., 2007; Sadasivam and DeCaprio, 2013; Sadasivam et al., 2012). LIN54 is crucial for the function of the MuvB core. It has been shown that mammalian mutants of LIN54 are arrested in the cell cycle at the G2 phase (Matsuo et al., 2012; Schmit et al., 2009). Moreover, CHC proteins such as LIN54 and MSL2 in animals bind DNA directly with their CHC domain in a sequence-specific manner (Marceau et al., 2016; Zheng et al., 2014).

In plants, a similar function of CHC proteins in cell cycle regulation has been observed. An Arabidopsis homolog of *LIN54*, *TSO1* (TSO means "ugly" in Chinese), was first reported to regulate cell division and organ differentiation in the shoot apical meristem. The inflorescence of *tso1-1* was unable to differentiate properly, and the resulting flower organization was severely impaired. Detailed studies of floral meristems showed that the cell walls of *tso1-1* were partially formed, and the nuclear DNA content was distributed more widely in *tso1-1* (2C-63C) than in wild-type (2C-4C), suggesting that the CHC proteins may be important for the mitosis of the meristem cells in Arabidopsis (Hauser et al., 2000; Liu et al., 1997; Sijacic et al., 2011; Song et al., 2000). Additionally, TSO1 has been shown to interact with MYB3R1, a plant transcription factor homologous to the mammalian BMYB factor, in different tissues to mediate cell differentiation of Arabidopsis (Wang et al., 2018). As mentioned previously, the interaction between CHC proteins and MYBs can also be found in the DREAM-like complex in Arabidopsis, which involves the CHC-containing TSO1-like CXC 5 (TCX5) protein and three MYB repressors.

Similar to animals, this DREAM-like complex is essential for repressing the S/M-specific genes during the G2 phase. The function of TCX5 on this complex remains unclear (Kobayashi et al., 2015).

Overview of the thesis work

In this thesis, I continue to investigate the role CHT7 plays in the cell cycle of Chlamydomonas. Previous studies of CHT7 have thoroughly characterized the impact of CHT7 on the transition between the N-replete cell division cycle and the N deprivation-induced quiescence state. Here, I extend the scope and ask whether the CHT7 protein is fundamentally involved in programming transcription throughout the cell cycle under various growth conditions. Chapter 2 presents a comprehensive transcriptome analysis of the *cht7* mutant under synchronous growth during N-replete, ND, and NR conditions, aiming to define the CHT7-mediated gene networks throughout the cell cycle. Chapter 3 examines the potential elements in the promoters of the misregulated genes in *cht7*, providing insights into the molecular mechanism of CHT7-mediated gene regulation. Chapter 4 explores the potential relationships of CHT7 to other cell cycle regulators, including a newly characterized CHC protein in Chlamydomonas. Finally, Chapter 5 summarizes the work above and provides directions for future study.

APPENDIX

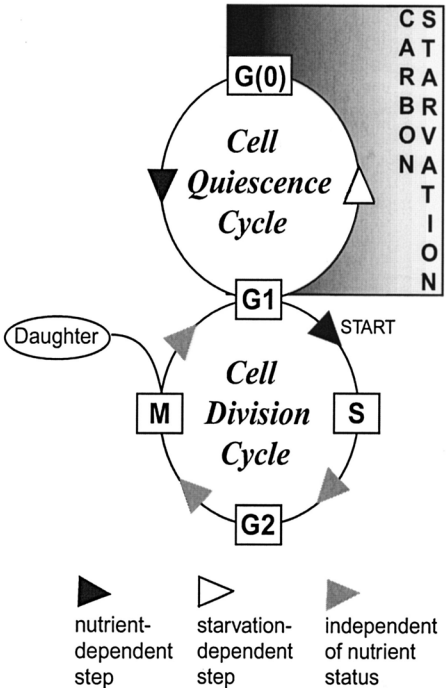


Figure 1.1. Metabolic state-dependent cell cycle progression.

A general overview of the cell cycle of microorganisms and the nutrient-dependence of each cell cycle stage (obtained from Gray et al., 2004).

Crop	Oil Yield (gal/acre/yr)
Soybean	48.0
Camelina	59.8
Sunflower	101.9
Jatropha	201.7
Oil palm	634.0
Algae*	1,500 (FY14) 2,500 (FY 18) 3,700 (FY20) 5,000 (FY22)

Source: Adapted from Darzins et al. (2010). *Note:* *Algae targets are set in the Bioenergy Technologies Office Multi-Year Program Plan (DOE 2016a) for intermediates.

Figure 1.2. Oil productivity of first-, second-, and third-generation biofuels.

A comparison of the oil yield (gal^{-1} $acre^{-1}$ $year^{-1}$) between different oil crops and projections of algal oil production (obtained from Barry et al., 2016).

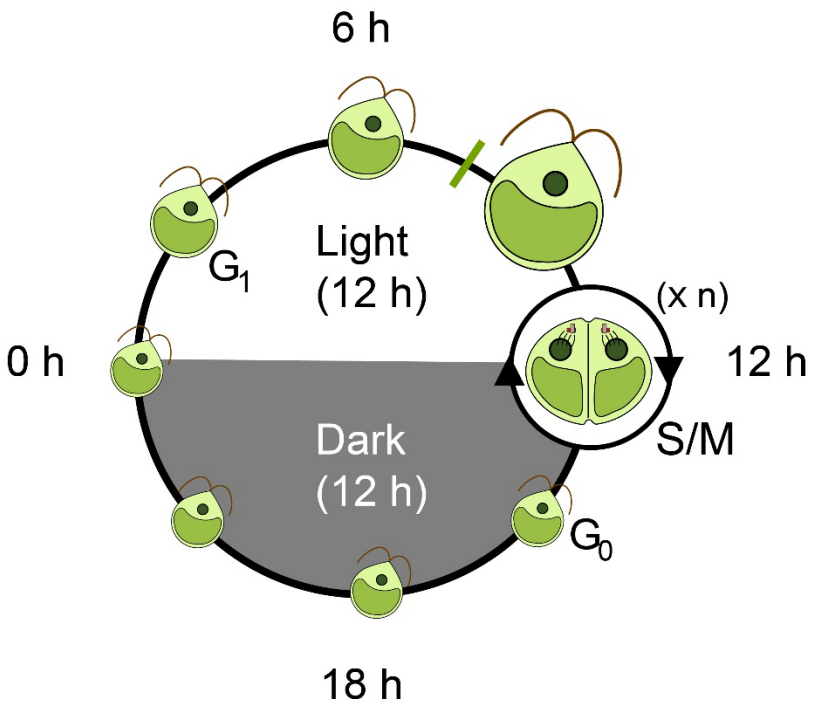


Figure 1.3. Cell cycle progression of Chlamydomonas during N-replete synchronous growth.

The diagram describes the progression of the cell cycle of wild-type cells under autotrophic diurnal conditions. Cells gain in size during the day (G1 phase). In the late G1 stage, cells passing the cell-size dependent commitment point (green bar) enter the S/M phase (S: DNA replication; M: mitosis) and undergo multiple rounds (n) of cell division. The new daughter cells stay in the G0 phase and start a new cycle the next day (0h) (modified from Lin et al., 2022).

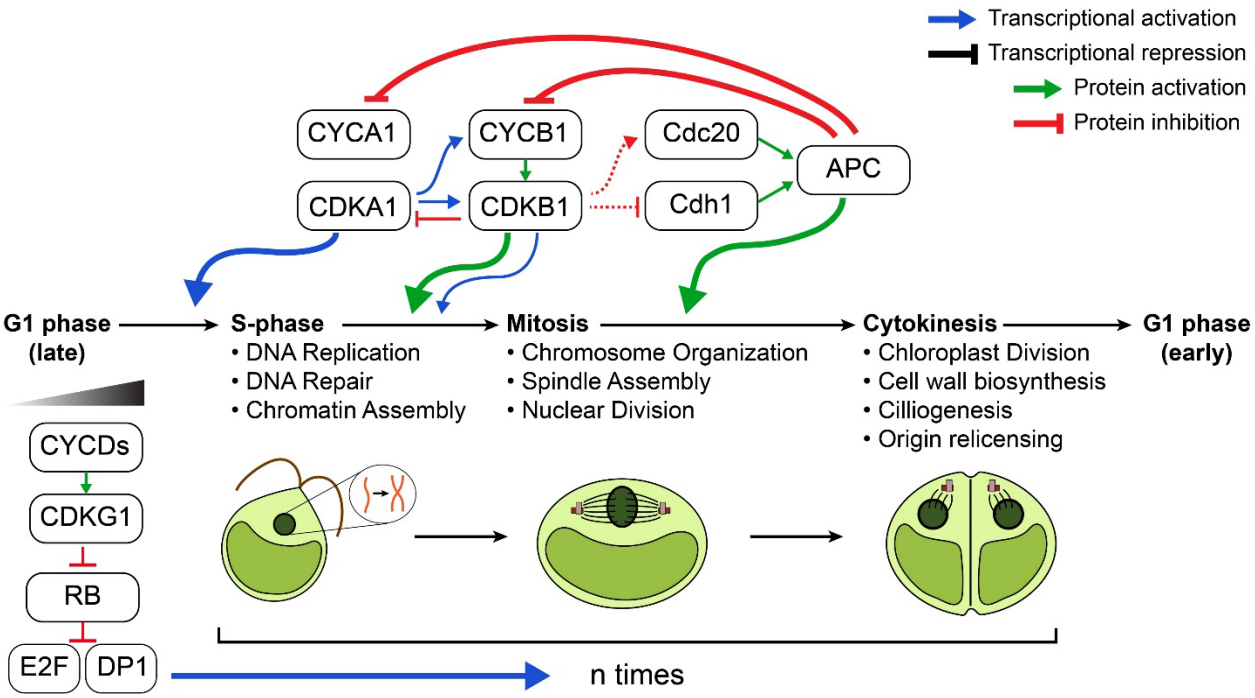


Figure 1.4. Regulation of the Chlamydomonas cell cycle.

Current understanding of the coordination between regulatory protein complexes of the cell cycle. RB: Retinoblastoma protein; E2F: E2 factor; DP1: Dimerization Partner 1; CYC: cyclin; CDK: Cyclin-Dependent Kinase; APC: Anaphase Promoting Complex; CDC20: Cell Division Cycle 20; CDH1: Cdc20 Homologue-1 (adapted from Lin et al., 2022).

REFERENCES

REFERENCES

Atkins, K.C., and Cross, F.R. (2018). Interregulation of CDKA/CDK1 and the plant-specific cyclin-dependent kinase CDKB in control of the Chlamydomonas cell cycle. Plant Cell 30:429-446.

Barry, A., Wolfe, A., English, C., Ruddick, C., and Lambert, D. (2016). 2016 National Algal Biofuels Technology Review. 2016 National Algal Biofuels Technology Review.

Beall, E.L., Manak, J.R., Zhou, S., Bell, M., Lipsick, J.S., and Botchan, M.R. (2002). Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. Nature 420:833-837.

Beck, C.F., and Acker, A. (1992). Gametic Differentiation of Chlamydomonas reinhardtii: Control by Nitrogen and Light. Plant Physiol 98:822-826.

Boyle, N.R., Page, M.D., Liu, B., Blaby, I.K., Casero, D., Kropat, J., Cokus, S.J., Hong-Hermesdorf, A., Shaw, J., Karpowicz, S.J., et al. (2012). Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in Chlamydomonas. J Biol Chem 287:15811-15825.

Cross, F.R. (2020). Regulation of multiple fission and cell-cycle-dependent gene expression by CDKA1 and the RB-E2F pathway in Chlamydomonas. Curr Biol 30:1855-1865 e1854.

Cross, F.R., and Umen, J.G. (2015). The Chlamydomonas cell cycle. Plant J 82:370-392.

Dastmalchi, M. (2021). Late-stage maturation erodes lipid content in soybean seeds. Plant Physiol 186:816-818.

Deshpande, A., Sicinski, P., and Hinds, P.W. (2005). Cyclins and cdks in development and cancer: a perspective. Oncogene 24:2909-2915.

Desvoyes, B., and Gutierrez, C. (2020). Roles of plant retinoblastoma protein: cell cycle and beyond. EMBO J 39:e105802.

Draetta, G. (1990). Cell-cycle control in eukaryotes - molecular mechanisms of Cdc2 activation. Trends in Biochemical Sciences 15:378-383.

Eloy, N.B., Lima, M.D., Ferreira, P.C.G., and Inze, D. (2015). The role of the anaphase-promoting complex/cyclosome in plant growth. Crit Rev Plant Sci 34:487-505.

Engel, B.E., Cress, W.D., and Santiago-Cardona, P.G. (2015). The retinoblastoma protein: A master tumor suppressor acts as a link between cell cycle and cell adhesion. Cell Health Cytoskelet 7:1-10.

Enserink, J.M., and Kolodner, R.D. (2010). An overview of Cdk1-controlled targets and processes. Cell Div 5:11.

Fang, S.C., de los Reyes, C., and Umen, J.G. (2006). Cell size checkpoint control by the retinoblastoma tumor suppressor pathway. PLoS Genet 2:e167.

Friend, S.H., Bernards, R., Rogelj, S., Weinberg, R.A., Rapaport, J.M., Albert, D.M., and Dryja, T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 323:643-646.

Gray, J.V., Petsko, G.A., Johnston, G.C., Ringe, D., Singer, R.A., and Werner-Washburne, M. (2004). "Sleeping beauty": quiescence in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev 68:187-206.

Guarnieri, M.T., Nag, A., Smolinski, S.L., Darzins, A., Seibert, M., and Pienkos, P.T. (2011). Examination of triacylglycerol biosynthetic pathways via de novo transcriptomic and proteomic analyses in an unsequenced microalga. PLoS One 6:e25851.

Hallenbeck, P.C., Grogger, M., Mraz, M., and Veverka, D. (2016). Solar biofuels production with microalgae. Appl Energ 179:136-145.

Hauser, B.A., He, J.Q., Park, S.O., and Gasser, C.S. (2000). TSO1 is a novel protein that modulates cytokinesis and cell expansion in Arabidopsis. Development 127:2219-2226.

Khan, M.I., Shin, J.H., and Kim, J.D. (2018). The promising future of microalgae: current status, challenges, and optimization of a sustainable and renewable industry for biofuels, feed, and other products. Microb Cell Fact 17:36.

Kobayashi, K., Suzuki, T., Iwata, E., Magyar, Z., Bogre, L., and Ito, M. (2015). MYB3Rs, plant homologs of Myb oncoproteins, control cell cycle-regulated transcription and form DREAM-like complexes. Transcription 6:106-111.

Li-Beisson, Y., Thelen, J.J., Fedosejevs, E., and Harwood, J.L. (2019). The lipid biochemistry of eukaryotic algae. Prog Lipid Res 74:31-68.

Li, Y., Liu, D., Lopez-Paz, C., Olson, B.J., and Umen, J.G. (2016). A new class of cyclin dependent kinase in Chlamydomonas is required for coupling cell size to cell division. Elife 5:e10767.

Lien, T., and Knutsen, G. (1979). Synchronous growth of *Chlamydomonas reinhardtii* (chlorophyceae) - review of optimal conditions. Journal of Phycology 15:191-200.

Litovchick, L., Sadasivam, S., Florens, L., Zhu, X., Swanson, S.K., Velmurugan, S., Chen, R., Washburn, M.P., Liu, X.S., and Decaprio, J.A. (2007). Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. Molecular Cell 26:539-551.

Liu, B., and Benning, C. (2013). Lipid metabolism in microalgae distinguishes itself. Curr Opin Biotechnol 24:300-309.

Liu, Z., Running, M.P., and Meyerowitz, E.M. (1997). TSO1 functions in cell division during Arabidopsis flower development. Development 124:665-672.

Lucker, B.F., Hall, C.C., Zegarac, R., and Kramer, D.M. (2014). The environmental photobioreactor (ePBR): An algal culturing platform for simulating dynamic natural environments. Algal Res 6:242-249.

Marceau, A.H., Felthousen, J.G., Goetsch, P.D., Iness, A.N., Lee, H.W., Tripathi, S.M., Strome, S., Litovchick, L., and Rubin, S.M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun 7:12301.

Matsuo, T., Kuramoto, H., Kumazaki, T., Mitsui, Y., and Takahashi, T. (2012). LIN54 harboring a mutation in CHC domain is localized to the cytoplasm and inhibits cell cycle progression. Cell Cycle 11:3227-3236.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., et al. (2010). Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol 154:1737-1752.

Nowack, M.K., Harashima, H., Dissmeyer, N., Zhao, X., Bouyer, D., Weimer, A.K., De Winter, F., Yang, F., and Schnittger, A. (2012). Genetic framework of cyclin-dependent kinase function in Arabidopsis. Dev Cell 22:1030-1040.

Olson, B.J., Oberholzer, M., Li, Y., Zones, J.M., Kohli, H.S., Bisova, K., Fang, S.C., Meisenhelder, J., Hunter, T., and Umen, J.G. (2010). Regulation of the Chlamydomonas cell cycle by a stable, chromatin-associated retinoblastoma tumor suppressor complex. Plant Cell 22:3331-3347.

Park, J.J., Wang, H., Gargouri, M., Deshpande, R.R., Skepper, J.N., Holguin, F.O., Juergens, M.T., Shachar-Hill, Y., Hicks, L.M., and Gang, D.R. (2015). The response of *Chlamydomonas reinhardtii* to nitrogen deprivation: a systems biology analysis. Plant J 81:611-624.

Pienkos, P.T., and Darzins, A. (2009). The promise and challenges of microalgal-derived biofuels. Biofuel Bioprod Bior 3:431-440.

Saad, M.G., Dosoky, N.S., Zoromba, M.S., and Shafik, H.M. (2019). Algal biofuels: Current status and key challenges. Energies 12.

Sadasivam, S., and DeCaprio, J.A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. Nat Rev Cancer 13:585-595.

Sadasivam, S., Duan, S., and DeCaprio, J.A. (2012). The MuvB complex sequentially recruits B-Myb and FoxM1 to promote mitotic gene expression. Genes Dev 26:474-489.

Salazar-Roa, M., and Malumbres, M. (2017). Fueling the Cell Division Cycle. Trends Cell Biol 27:69-81.

Satyanarayana, K.G., Mariano, A.B., and Vargas, J.V.C. (2011). A review on microalgae, a versatile source for sustainable energy and materials. Int J Energ Res 35:291-311.

Schmit, F., Cremer, S., and Gaubatz, S. (2009). LIN54 is an essential core subunit of the DREAM/LINC complex that binds to the cdc2 promoter in a sequence-specific manner. FEBS J 276:5703-5716.

Schmollinger, S., Muhlhaus, T., Boyle, N.R., Blaby, I.K., Casero, D., Mettler, T., Moseley, J.L., Kropat, J., Sommer, F., Strenkert, D., et al. (2014). Nitrogen-sparing mechanisms in Clamydomonas affect the transcriptome, the proteome, and photosynthetic metabolism. Plant Cell 26:1410-1435.

Sijacic, P., Wang, W., and Liu, Z. (2011). Recessive antimorphic alleles overcome functionally redundant loci to reveal TSO1 function in Arabidopsis flowers and meristems. PLoS Genet 7:e1002352.

Song, J.Y., Leung, T., Ehler, L.K., Wang, C., and Liu, Z. (2000). Regulation of meristem organization and cell division by TSO1, an Arabidopsis gene with cysteine-rich repeats. Development 127:2207-2217.

Strenkert, D., Schmollinger, S., Gallaher, S.D., Salome, P.A., Purvine, S.O., Nicora, C.D., Mettler-Altmann, T., Soubeyrand, E., Weber, A.P.M., Lipton, M.S., et al. (2019). Multiomics resolution of molecular events during a day in the life of Chlamydomonas. Proc Natl Acad Sci U S A 116:2374-2383.

Takeuchi, T., Sears, B.B., Lindeboom, C., Lin, Y.T., Fekaris, N., Zienkiewicz, K., Zienkiewicz, A., Poliner, E., and Benning, C. (2020). Chlamydomonas CHT7 is required for an effective quiescent state by regulating nutrient-responsive cell cycle gene expression. Plant Cell 32:1240-1269.

Thomas, J.H., Ceol, C.J., Schwartz, H.T., and Horvitz, H.R. (2003). New genes that interact with lin-35 Rb to negatively regulate the let-60 ras pathway in *Caenorhabditis elegans*. Genetics 164:135-151.

Tsai, C.H., Uygun, S., Roston, R., Shiu, S.H., and Benning, C. (2018). Recovery from N deprivation is a transcriptionally and functionally distinct state in Chlamydomonas. Plant Physiol 176:2007-2023.

Tsai, C.H., Warakanont, J., Takeuchi, T., Sears, B.B., Moellering, E.R., and Benning, C. (2014). The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in Chlamydomonas. Proc Natl Acad Sci U S A 111:15833-15838.

Tulin, F., and Cross, F.R. (2014). A microbial avenue to cell cycle control in the plant superkingdom. Plant Cell 26:4019-4038.

Tulin, F., and Cross, F.R. (2015). Cyclin-dependent kinase regulation of diurnal transcription in Chlamydomonas. Plant Cell 27:2727-2742.

Umen, J.G. (2005). The elusive sizer. Curr Opin Cell Biol 17:435-441.

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

Valcourt, J.R., Lemons, J.M., Haley, E.M., Kojima, M., Demuren, O.O., and Coller, H.A. (2012). Staying alive: metabolic adaptations to quiescence. Cell Cycle 11:1680-1696.

Vallon, O., and Wollman, F.A. (1995). Mutations Affecting O-Glycosylation in Chlamydomonas reinhardtii Cause Delayed Cell Wall Degradation and Sex-Limited Sterility. Plant Physiol 108:703-712.

Wang, W., Sijacic, P., Xu, P., Lian, H., and Liu, Z. (2018). Arabidopsis TSO1 and MYB3R1 form a regulatory module to coordinate cell proliferation with differentiation in shoot and root. Proc Natl Acad Sci U S A 115:E3045-E3054.

Zheng, S., Villa, R., Wang, J., Feng, Y., Wang, J., Becker, P.B., and Ye, K. (2014). Structural basis of X chromosome DNA recognition by the MSL2 CXC domain during Drosophila dosage compensation. Genes Dev 28:2652-2662.

Zones, J.M., Blaby, I.K., Merchant, S.S., and Umen, J.G. (2015). High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. Plant Cell 27:2743-2769.

CHAPTER 2
Chlamydomonas CHT7 is involved in repressing DNA replication and mitotic genes during synchronous growth[1]

[1]This work has been published with minor edits in G3: Genes | Genomes | Genetics.

Lin, Y.T., Takeuchi, T., Youk, B., Umen, J., Sears, B.B., and Benning, C. (2022). Chlamydomonas CHT7 is involved in repressing DNA replication and mitotic genes during synchronous growth. G3 Genes | Genomes | Genetics, 12. https://doi.org/10.1093/g3journal/jkac023

AUTHOR CONTRIBUTIONS

T.T. and B.B.S. developed strains. Y.-T.L. and T.T. designed experiments. T.T. and B.Y. assisted Y.-T.L. in conducting experiments. Y.-T.L. analyzed data. C.B. designed and coordinated the study. Y.-T.L. drafted the article which was edited by T.T., J.U., B.B.S and C.B.

ABSTRACT

In the green alga *Chlamydomonas reinhardtii*, regulation of the cell cycle in response to external cues is critical for survival in a changing environment. The loss of the nuclear COMPROMISED HYDROLYSIS OF TRIACYLGLYCEROLS 7 (CHT7) protein affects the expression of many genes especially in response to nitrogen availability. Cells lacking CHT7 exhibit abnormal cell morphology following nitrogen deprivation (ND) and fail to resume normal cell division after N resupply. To investigate the function of CHT7 in the regulation of cell cycle-related pathways, cells were synchronized, and RNA-seq analysis was performed during various stages of the cell cycle. In the *cht7* mutant following ND, the cells were not dividing, but a subset of cell cycle genes involved in DNA replication and mitosis were found to be derepressed, suggesting that the CHT7 protein plays a role in cell cycle regulation that is opposite to that of the mitotic cyclin-dependent kinases. Furthermore, genes for cell wall synthesis and remodeling were found to be abnormally induced in non-dividing *cht7* cells; this misregulation may deplete cellular resources and thus contribute to cell death following ND. Lastly, 43 minimally characterized kinases were found to be highly misregulated in *cht7*. Further analysis suggested that some of these CHT7-regulated kinases may be related to the MAP3K and Aurora-like kinases, while others are unique. Together, these results suggest a role of CHT7 in transcriptional regulation of the cell cycle and reveal several pathways and genes whose expression appears to be subject to a CHT7-mediated regulatory network.

INTRODUCTION

In the unicellular green alga Chlamydomonas (*Chlamydomonas reinhardtii*), orderly progression through the appropriate stages of the cell cycle is critical for survival in a changing environment. Different stages of the cell cycle exhibit specific metabolic activities and distinct cytological phases of cell division. Cellular metabolism and cell division are tightly coordinated to optimize the utilization of energy for cell growth and proliferation (Salazar-Roa and Malumbres, 2017). The regulation of the cell cycle is modulated by the programming of gene expression. An RNA sequencing (RNA-seq) analysis by Zones et al. (2015) has documented the transcriptome dynamics throughout the life cycle of autotrophically growing wild-type Chlamydomonas cells under diel cycles. Briefly, genes of photosynthesis are highly expressed in the G1 stage during the daytime to ensure rapid cell growth driven by light energy. On the other hand, genes involved in cell division are not expressed until the end of the day, after the cell size-determined commitment point, when cells enter the S/M phase (S: DNA synthesis; M: mitosis). The peak expression of DNA replication, nuclear division, and chloroplast fission genes at this stage indicates rapid cycles of cell division. Cells that finish dividing then enter a short G0 period in the dark, with downregulation of cell division and cellular metabolism genes, while genes of flagellar assembly are actively expressed, after which the daughter cells will hatch out at dawn to begin a new life cycle (Zones et al., 2015). Together, the oscillating gene expression of cell division and metabolic pathways indicates the importance of transcriptional regulation during progression through the cell cycle.

In addition to the normal cell division cycle, transcriptional cell cycle control has also been observed during the transition into quiescence, a prolonged G0 phase, under nutrient-limited conditions (Gray et al., 2004). For example, Chlamydomonas following N deprivation has been

reported to down-regulate the transcription of genes encoding proteins that participate in photosynthesis, aerobic respiration, protein translation, and cell division. On the other hand, genes involved in triacylglycerol (TAG) and starch biosynthesis are upregulated in N-deprived cells, allowing the accumulation of carbon storage compounds (Miller et al., 2010; Park et al., 2015; Schmollinger et al., 2014). These transcriptional controls are required for the proper entry into quiescence and survival under stress.

The identification of COMPROMISED HYDROLYSIS OF TRIACYLGLYCEROLS 7 (CHT7, Cre11.g481800) has provided insight into the regulatory mechanism of quiescence transition in Chlamydomonas. The CHT7 loss-of-function mutant, *cht7*, was isolated during a screen for Chlamydomonas mutants that were slow or unable to degrade TAG when N is resupplied after N deprivation. Degradation of TAG allows the cell to reutilize the carbon and energy stored in that molecule. RNA-seq analysis of non-synchronized, mixotrophic cells during the same study suggested that CHT7 is required for cells to transition between quiescence and the cell division cycle (Tsai et al., 2014). Subsequent work has shown that the expression of some cell cycle genes was strongly derepressed in *cht7* cells following N deprivation (Takeuchi et al., 2020b). A fraction of the N-deprived *cht7* cells were observed to have multiple nuclei and a swollen cell appearance, which was not seen in the *cht7* cells under N-replete (N+) conditions although the cell cycle genes were slightly misregulated (Takeuchi et al., 2020b). Most of the N-deprived *cht7* cells were unresponsive to N-refeeding, leading to the overall delayed TAG degradation mentioned above (Takeuchi et al., 2020b). Together, these results strongly suggested that CHT7 is required for viability during quiescence and for the proper transition from quiescence back into the normal cell division cycle.

The CHT7 protein has been localized to the nucleus and contains a CHC (CXC-Hinge-CXC) domain that is composed of two cysteine-rich CXC motifs (Tsai et al., 2014). Previous studies have shown that CHC proteins in plants and animals play an essential role in cell cycle regulation by forming complexes with various transcription factors and binding to DNA directly with their CHC domain (Beall et al., 2002; Marceau et al., 2016; Matsuo et al., 2012; Sadasivam and DeCaprio, 2013; Schmit et al., 2009; Sijacic et al., 2011; Thomas et al., 2003; Wang et al., 2018; Zheng et al., 2014). However, it remains unclear whether CHT7 in Chlamydomonas functions through the same molecular mechanism. Our recent studies (Takeuchi et al., 2020b) suggested that the predicted DNA binding domain of the CHT7 protein is not essential or is redundant as its deletion did not cause a detectable phenotype. In contrast, removal of the C-terminal disordered region resulted in a *cht7* null phenotype. This finding implicated a functional mechanism for CHT7 that may be distinct from other known CHC proteins and that does not require direct DNA binding by CHT7 through its CHC domain. Until now most CHT7 studies have focused on the transition of quiescence following N deprivation and N resupply; the role of CHT7 during the normal cell division cycle has received less attention. One reason is that previous studies were done using asynchronous cell cultures (Takeuchi et al., 2020b; Tsai et al., 2014). To address this issue, we recently used cell cultures grown under conditions optimized for synchronization of cell division to investigate the impact of CHT7 on cell cycle progression, and we found that CHT7 proteins are more abundant during the S/M phase and may interact with the RB protein, suggesting that CHT7 might be a part of the cell cycle regulatory network (Takeuchi et al., 2020a). Here, to better understand the role of CHT7 in cell cycle regulation and to globally characterize the effect of CHT7 on gene expression under various conditions, we performed an RNA-seq of *cht7* and two control strains grown under light-dark conditions to achieve

synchronized division. Our data show that genes involved in DNA replication, mitosis, and cell cycle regulation were inappropriately derepressed in *cht7* during quiescence and the G0 and G1 phases of the normal cell division cycle. We also found abnormally induced expression of genes encoding cell-wall components and potential regulators of cell wall remodeling in *cht7* following N deprivation. Lastly, we identified a previously uncharacterized group of protein kinases whose gene expression was subject to regulation by CHT7.

RESULTS

Cell cycle analysis under synchronizing conditions

Our previous observations showed largely synchronized growth of the *cht7* mutant and wild-type cells in photobioreactors in the presence of N while the cultures were bubbled with CO_2 (Takeuchi et al., 2020a). To investigate the role of CHT7 in cell cycle regulation, three strains, wild type (WT), the *cht7* mutant, and the *cht7* complemented line *CHT7-HA::cht7*, were cultured under three sequential growth conditions determined by the availability of N: N-replete (N+), N-deprived (ND), and N resupplied (NR). N+ was defined as the standard growth condition, and samples harvested from cultures grown under this condition were collected at different time points ZTn, where n represents the time after the onset of light (ZT stands for Zeitgeber time, e.g., ZT6 was taken 6 h after illumination). Similarly, the NDn samples were taken n hours after N deprivation; these should represent cells in the quiescent state. The NRn samples were taken n hours following N resupply to investigate the transition from quiescence to the normal cell division cycle. We observed the typical progression of the cell cycle and the cellular response to N+, ND, and NR with cell size distribution analysis (Supplemental Fig. S2.1, left column). In all instances the cells increased size from ZT0 to ZT10 and underwent cell division between ZT12 and ZT16. Cells

finished dividing and gave rise to single, small cells at ZT24 (ZT0) entering the next G1 phase. To obtain an independent quantitative measure for cell cycle progression and synchronization, we determined the DNA content of individual cells (Supplemental Fig. S2.1, right column). Our previous microscopy studies of DAPI-stained cells indicated that some N-deprived *cht7* cells are multi-nucleated (Takeuchi et al., 2020b; Tsai et al., 2014); this phenotype was not observed in *cht7* cells grown under standard N+ conditions. Here, we applied a more quantitative method using flow cytometry of cells stained with the DNA dye SYTOX^{TM} Green (Fig. S2.1, right column and Fig. 2.1A-B) following established protocols (Fang et al., 2006; Tulin and Cross, 2014). As shown in Fig. 2.1A, an average of ~96% of WT and *CHT7-HA::cht7* cells stained with SYTOX^{TM} Green and N+ grown were found to contain a haploid DNA content (1C) at ZT0-10, prior to DNA replication. In contrast, only 80% of the stained *cht7* cells had a 1C DNA content with almost 20% carrying two or more genomes (≥2C) at the same timepoints (Fig. 2.1A; also see Fig. 2.1B for an example at ZT6). In addition, there appeared to be more cell debris (Fig. 2.1B, non-cells) and cells with irregular DNA contents (Fig. 2.1B, UD-S and UD-L) across all N+, ND, and NR conditions for the *cht7* cultures. While the loss of viability of *cht7* cells under ND and NR has been previously reported using different methods (Takeuchi et al., 2020b), these more sensitive flow cytometry results point to some loss of viability under regular N+ growth conditions. Finally, at timepoints of cell division under N+ growth (S/M, Fig. 2.1A), the population of cells containing ≥2C DNA increased far more in *cht7* (30.5%) than in the WT and *CHT7-HA::cht7* line (21.2% and 13.4% respectively). Furthermore, under ND and NR culture conditions, *cht7* cells contained a ≥ 2C DNA content in a larger fraction of cells than did WT and *CHT7-HA::cht7* cultures (Fig. 2.1A and Fig. S2.1). These findings were consistent with our previous observations of N-deprived *cht7* cells (Takeuchi et al., 2020b). The persistent presence of a significant fraction of *cht7* cells with ≥2C

chromatin content throughout the sampling regime suggests that in the WT, the CHT7 protein may suppress the initiation of DNA replication. Another possibility is that cells lacking CHT7 may be impaired in cytokinesis and thus carry ≥2C DNA. Impairment of these processes in the *cht7* mutant could impede complete synchronization of *cht7* cultures, as discussed further below.

To observe the impact of the loss of CHT7 on the G1 phase under N+ conditions, we examined *cht7* cells collected at ZT6 by microscopy. As shown in Fig. 2.1C and 2.1D, cells were classified based on their morphology and viability. The majority of WT and *CHT7-HA::cht7* cells were single and flagellated (type 1, 89.2-91.5%). In the *cht7* mutant, there were more cells with diverse morphological phenotypes such as round and non-flagellated cells (type 2, 19.3-22.9%), dividing cells (type 3, 1.7-3.4%), others (type 4, 8.4-10.0%), and dead cells (type 5, 4.8-5.7%). Together, these results indicated that *cht7* cells have a broader range of nuclear contents and cell types than do the two wild-type controls under conditions that synchronize the cell cultures, possibly due to an altered regulation of the cell cycle in the mutant.

Transcriptome dynamics of cell cultures under various conditions

Because the decreased synchrony and increased DNA content of *cht7* cultures point towards a possible role of CHT7 in cell cycle regulation of Chlamydomonas, we set out to better understand the scope of the transcriptional networks that could be affected by CHT7, and we performed RNA-seq analysis of WT, *CHT7-HA::cht7*, and *cht7* cell cultures. Samples were collected in biologically independent triplicates at nine representative timepoints under the N+, ND, and NR conditions corresponding to key points of the cell cycle and transitions in and out of quiescence (Fig. 2.2A).

After processing and mapping raw reads to the Chlamydomonas predicted transcriptome v5.6 (Creinhardtii_281_v5.6, Phytozome), we performed principal component analysis (PCA) to

compare the transcriptomes of samples in the N+, ND, and NR conditions. As shown in Figure 2.2B, samples with different transcriptomes form distinct groups across the dimensions of PC1 and PC2, the two principal components contributing most to the total variance between the samples in each condition. Samples with similar transcriptome compositions, on the other hand, were close to each other (e.g. Wild type and *CHT7-HA::cht7*). The transcriptomes of *cht7* from the ND (middle plot; Fig. 2.2B), NR (lower plot; Fig. 2.2B), and N+ (ZT6) (circle dots; upper plot; Fig. 2.2B) samplings were most different from the respective two control samples, suggesting specific functions of CHT7 at these timepoints. Notably, under N+ conditions, the ZT0 samples also compose a unique group on the plot, because they represent the only time points at which cells were harvested in the dark (diamond dots; upper plot; Fig. 2.2B). As expected, light has a strong impact on the transcriptome of Chlamydomonas. It should be noted that *cht7* samples for ZT0 fell into the same group as WT and *CHT7-HA::cht7* suggesting that light-responsive signaling is functional in the mutant, which is a prerequisite for synchronizing cell division.

Next, we performed hierarchical gene clustering to group the transcription profiles of 14,946 transcripts across all samples with an average read count > 20 (n=3), settling on 14 gene clusters with unique expression profiles (C01-14; Fig. 2.2C). Gene Ontology (GO) enrichment analysis showed that each gene cluster was annotated with specific functions correlated to their expression pattern (Fig. 2.2D). For example, genes for cilium assembly and cytoplasmic translation were found in the C01 and C02 clusters, respectively, which were mainly transcribed in the dark in all strains, as reported previously (Zones et al., 2015). On the other hand, GO terms related to photosynthetic processes were highly enriched in group C05 genes, which were mainly transcribed at ZT6 in all strains, indicating that cells at this timepoint were highly active photosynthetically. Meanwhile, genes falling into the C11 and C12 clusters were highly expressed

in ND samples of all strains, and GO terms enriched in these two clusters were mostly related to the metabolism of N-containing and other carbon compounds as well as stress responses, suggesting an adjustment of metabolic pathways in response to N deprivation (Fig. 2.2C and 2.2D). In summary, the transcriptional profiles of metabolic and cell cycle genes of the two control Chlamydomonas strains were generally consistent with previous transcriptomic studies of synchronized N+ cells (Zones et al., 2015) and ND cells (Miller et al., 2010; Schmollinger et al., 2014), allowing us to focus on the differentially expressed genes in the *cht7* mutant.

A global analysis of upregulated genes in the *cht7* mutant

We performed differential expression (DE) analysis to compare the read count of all individual transcripts (n=19,923) of any two strains at the same timepoints (e.g., *cht7* vs. WT at ZT6). The resulting differentially expressed genes (DEGs; $\log_2$ fold change >1 [up] or <1 [down] and padj < 0.05) are displayed in Figure 2.3A. Notably, when comparing *cht7* to the WT or the complemented line, the upregulated DEGs clearly outnumber the downregulated DEGs. This pattern was found at all time points except ZT0 (Fig. 2.3A), at which ~3,700 genes were down-regulated in both *cht7* and the complemented line, *CHT7HA::cht7*, compared to WT. This difference might be due to the deletion of several other genes in the background of the *cht7* mutant (Tsai et al., 2014). GO enrichment analysis showed that many of the specifically downregulated genes in *cht7* and *CHT7HA::cht7* versus the WT participate in processes such as cilium assembly and cell motility (data not shown). Because these downregulated DEGs at ZT0 are likely attributable to the original *cht7* mutant background, we will not discuss them further in this study, but focus on DEGs directly related to the loss of CHT7.

To identify differential expression that is specifically caused by the loss of CHT7 but not the other gene deletions in the *cht7* background, we compared the three groups at each timepoint (Fig. 2.3B). To be designated as a *cht7*-specific DEG, the gene must be differentially expressed in *cht7* relative to the controls and not be differentially expressed between *CHT7-HA::cht7* and WT. With this approach, we identified *cht7*-specific DEGs that were either up- or downregulated in the absence of CHT7 (Fig. 2.3C). It was quite striking that numbers of upregulated DEGs were 2.9-14.5-fold higher than downregulated DEGs across all timepoints, indicating that CHT7 might primarily be involved in transcriptional repression (Fig. 2.3C).

To gain insights into the function of the DEGs due to the loss of CHT7, we performed GO enrichment analysis of the total 3,551 DEGs that were specifically upregulated in *cht7* at more than one timepoint; 184 GO terms of biological processes were found to be highly enriched ($p < 0.05$) in these genes (Supplemental Fig. S2.2). Among these GO terms, 35 were enriched in at least six timepoints and were, therefore, chosen for further study (Fig. 2.3D).

CHT7 suppresses transcription of genes involved in DNA replication and mitosis

A group of 23 GO terms (Fig. 2.3D; yellow) drew our attention because they were all related to processes during the S/M phase of the cell cycle (e.g., DNA replication, nuclear division, chloroplast fission) and were highly enriched in *cht7*-specific upregulated DEGs ($p<0.01$), suggesting that they are repressed in the presence of CHT7. Furthermore, this specific enrichment pattern supported the finding that genes corresponding to those cell cycle processes in the GO annotations were programmed to be transcriptionally activated during cell division ZT10, ZT12, and NR12 (Fig 2C; C07).

Derepressed cell cycle genes in *cht7* under both N+ and ND conditions have also been observed in previous studies using non-synchronized cell cultures (Takeuchi et al., 2020b; Tsai et al., 2014). To further specify the S/M processes potentially regulated by CHT7, we carefully investigated the annotated functions of the 512 genes under 23 GO terms mentioned above, among which 236 (46%) were upregulated in *cht7* at more than one timepoint. Information from multiple functional prediction databases and a previous study of cell cycle genes of Chlamydomonas (Zones et al., 2015) were included for this purpose.

The results showed that 116 of the 236 DEGs were associated with DNA replication and DNA-replication-coupled homologous recombination (HR) repair, interstrand crosslink (ICL) repair, and chromatin assembly (Fig. 2.4A). These processes are referred to as S-phase events. During the normal cell division cycle, most genes that participate in S-phase events are transcribed at ZT10 and ZT12 (S/M phase) and turned off at ZT0 (G0 phase) and ZT6 (G1 phase). During ND-induced quiescence, these genes are repressed and not expressed until 12 hours after N-resupply (Fig. 2.4A; see columns of WT and *CHT7-HA::cht7*). However, in the *cht7* mutant, many of these S-phase genes are not repressed, although the expression of these genes is not further increased at ZT10, 12, and NR12 either (Fig. 2.4A and 2.4B). Figure 2.4B shows the fluctuation of transcript levels of genes known to be involved in specific events of the S-phase such as pre-Replicative Complex assembly (ORC1, Cdt1, and MCM8), pre-Initiation Complex assembly (Cdc45 and GINS4), DNA synthesis (POLA2,RFA1, and TOP2), DNA repair (CTIP-related, DNA2-related, BRCA2-related, and RAD51), and nucleosome assembly (TOUSLED-LIKE, ASF1, HTV1, and HLM8) (De Benedetti, 2012; Fabry et al., 1995; Godin et al., 2016; Horard et al., 2018; MacNeill, 2010; Makharashvili and Paull, 2015; Nimonkar et al., 2011; Riera et al., 2017; Vignard et al., 2007; Zones et al., 2015) in the wild-type lines, in contrast to the high background

levels of expression in *cht7*. The constitutive expression of genes in *cht7* that are normally activated specifically in S-phase may result in the presence of ≥2C DNA cells in *cht7* cultures (Fig. 2.1). To test the possibility that replicative stress in the *cht7* mutant caused the observed phenotypes, we examined the expression of homologs of genes that encode proteins known to be involved in sensing or repairing DNA damage in plants such as ATR kinases ATR1, ATR2, ATR-like Cre13.g564350 and the CDK inhibitory kinase WEE1 (Fig. 2.4C) (Bisova et al., 2005; Nisa et al., 2019). There was no significant difference in the expression of these genes between *cht7* and the control strains, except for the gene encoding WEE1 which was abnormally expressed in a pattern similar to other derepressed S-phase genes in *cht7*.

Lastly, we found that 68 of the 104 genes participating in mitotic processes (chromosome condensation, sister chromatid cohesion, spindle assembly, nuclear division) and chloroplast division seemed to be constitutively expressed in *cht7* (Fig. 2.5A). Some of these genes are known to play a crucial role in the mitotic events following DNA replication in vertebrates and yeast (Chan et al., 2013; Cheeseman, 2014; Hirano, 2012; Musacchio and Salmon, 2007; Peters et al., 2008; Sundin et al., 2011; Tian and Kong, 2019) and chloroplast fission in plants (Chen et al., 2018; Zones et al., 2015), and their transcript levels are shown in detail in Figure 2.5B. However, it remains unclear whether mitosis of *cht7* cells was directly affected given that previous studies did not show obvious abnormal cell division phenotypes in the *cht7* mutant under N+ conditions (Takeuchi et al., 2020a; Takeuchi et al., 2020b). Together, these findings suggest that CHT7 is required for down-regulating the transcription of genes involved in S/M phase-related processes during both the regular cell division cycle and quiescence, and that in *cht7* DNA replication seems to be activated across all conditions.

How does CHT7 fit into the regulatory network of the cell cycle in Chlamydomonas?

The finding of derepressed S/M phase genes in *cht7* during the normal cell division cycle in N+-grown cells suggested that CHT7 might have a function in the regulatory network of the cell cycle in Chlamydomonas. It also raised the possibility that CHT7 contributes to modulating the expression of cell cycle regulators, which in turn affect the transcription of downstream targets. To test this hypothesis, we examined the expression of genes encoding known cell cycle regulators of Chlamydomonas (Fig. 2.6). Fig. 2.6A describes the current understanding of cell cycle regulation of Chlamydomonas. The timing of entry into the S/M phase is mainly controlled by a cyclin-dependent kinase, CDKA1, in a cell-size-dependent manner (Cross, 2020). In the late G1 phase, cells reaching the cell size threshold for cell division activate CDKA1 which induces the transcription of *CYCB1*, *CDKB1*, and other essential genes for DNA replication and mitosis (Atkins and Cross, 2018; Cross, 2020; Tulin and Cross, 2015). The initiation of the S/M program by CDKA1 triggers the activation of CDKB1 by CYCB1, which then inactivates CDKA1 (negative feedback) and promotes the entry into mitosis (e.g., spindle formation), inducing the re-initiation of DNA replication after the first round of nuclear division (Atkins and Cross, 2018; Tulin and Cross, 2015). Another important regulator for the progression of the cell cycle is the anaphase-promoting complex (APC), a universally conserved complex required for chromosome segregation (Cross and Umen, 2015). In Chlamydomonas, APC is required for both mitotic exit and the inactivation of CDKA1 and CYCB1-CDKB1 to ensure the transition into cytokinesis after the mitotic spindle is formed (Atkins and Cross, 2018). In addition to the induction of cell division above, the control of cell division activity during the S/M phase is an important aspect of cell cycle regulation. This pathway is governed by the Retinoblastoma complex (RB)-DP1-E2F which does not control timing of S/M but is required for appropriate coupling of cell division number and cell

cycle transitions to cell size. The D-cyclin dependent RB kinase CDKG1 is limiting for cell division number and is produced just prior to S/M phase with an abundance that positively correlates with mother cell-size. CDKG1-D cyclin along with other S/M kinases phosphorylate RB, inactivating its repressor function, and enabling the positive regulators of cell division, chromatin-bound heterodimeric DP and E2F subunits, to activate cell division (Cross, 2020; Li et al., 2016; Olson et al., 2010). This mechanism ensures that larger mother cells will divide more times than smaller mother cells and allows size homeostasis to be maintained. Notably, unlike CDKA1, the RB-DP1-E2F complex does not control the timing of cell division nor the transcription of most S/M genes, but rather participates in cell size regulation (Cross, 2020; Fang et al., 2006).

As reported previously, the expression of genes encoding CDKA1, CYCB1, CDKB1, CYCA1, and CYCAB1 peaks at ZT10-12 of WT cells (Zones et al., 2015). We confirmed this expression pattern for the wild-type lines, as well as for *cht7* (Fig. 2.6B). This result indicates that CHT7 does not act upstream of CDKA1 to regulate the S/M program, although some of the regulatory targets of both proteins seemed to overlap. However, our data also showed that genes for CYCA1, CYCB1, and CDKB1 were derepressed in *cht7* at non-cell dividing timepoints such as ZT0, ND6-36, and NR6 (Fig. 2.6B). This pattern is similar to other CHT7-regulated S/M phase genes (Fig. 2.4 and 2.5), implying that the abnormal S/M program in *cht7* might be caused by the misregulated expression of *CYCB1* and *CDKB1* among other relevant genes. Interestingly, the expression of most of the genes encoding APC components was only slightly affected in *cht7* (Fig. 2.6C), indicating these genes may be subject to a different regulatory pathway than CYCB1 and CDKB1. The exceptions were the genes encoding the APC activators CDC20 and CDC20 homologue 1 (CDH1) (Cross and Umen, 2015), which were highly induced in the *cht7* mutant (Fig

6C), indicating that APC activity might still be affected in the *cht7* mutant. Conceivably, an imbalance in expression of cell cycle regulatory genes may be responsible for the observed phenotypes. Lastly, our results showed that there was no impact of the *cht7* mutation on the transcript levels of genes encoding the RB-DP-E2F complex- or CYCD-CDKG1 (Fig. 2.6D and 2.6E). However, we cannot exclude the possibility that the protein activity of CHT7 is directly affected by the RB-mediated pathway given that a subfraction of CHT7 has been shown to associate with RB during the cell division cycle (Takeuchi et al., 2020a).

Nitrogen starvation-induced cell wall synthesis in the *cht7* mutant

In addition to cell division-related processes, we observed that the term GO:006075, defined as (1-3) Beta-D-Glucan Biosynthetic Process, was also enriched in the *cht7*-specific upregulated DEGs (Fig. 2.3D). Further analysis showed that 7 of the 11 annotated genes under this GO term were upregulated in *cht7* at more than three timepoints across all conditions. The GSL1 gene (Cre04.g214650; Callose synthase gene) was one of the DEGs that was highly transcribed in *cht7* at all timepoints, especially after ND12 and during N refeeding (Fig. 2.7C). Callose is a component of the cell wall in seed plants, and GSL8 in Arabidopsis was found to be essential for cell plate formation during cytokinesis (Chen and Kim, 2009; Chen et al., 2009). Although cell walls are structurally different in Chlamydomonas and seed plants, callose is present at least in the primary wall of zygotes of *Chlamydomonas monoica* (Salanga and Van Winkle-Swift, 2020). Hence the misregulation of the presumed GSL1 gene in *cht7* suggested that CHT7 may also affect aspects of cell wall biosynthesis in Chlamydomonas at certain stages of the cell cycle. To test this hypothesis, we examined the *cht7*-upregulated DEGs for other genes that might encode proteins with roles in cell wall-related processes. A group of genes encoding hydroxyproline-rich glycoproteins

(HRGPs), the major cell wall component of Chlamydomonas (Cronmiller et al., 2019), fell into this category. Many of the *cht7*-induced HRGP genes belong to the Pherophorin (PHC) subfamily of HRGPs that was originally named for its C-terminal similarity to the sex-inducing pheromone of *Volvox carteri* (Sumper et al., 1993; von der Heyde and Hallmann, 2020). We identified 109 genes potentially encoding cell-wall-associated HRGPs, including 67 PHCs, in the Chlamydomonas genome v5.6, and 68 of the respective genes (including those for 45 pherophorins) were upregulated in *cht7* at more than three timepoints. The Z-score heatmap of these cell wall protein-encoding genes revealed that most of them were highly induced in *cht7* after 12 hours of N deprivation (Fig. 2.7A and 2.7B).

To better understand how CHT7 could modulate the expression of genes encoding PHCs and other HRGPs, we considered two instances of cell wall biosynthesis in Chlamydomonas that have been previously reported. The first one is the biosynthesis of new daughter cell walls following nuclear division. A transcriptomic study reported that the increase in expression of some PHC encoding genes was highly correlated with the S/M phase of the cell cycle and dependent on the kinases CDKA1 and CDKB1(Tulin and Cross, 2015). For example, the gene for GAS28, a cell wall HRGP, was not transcribed in the temperature-arrested *cdkb1* mutant (Cronmiller et al., 2019; Hoffmann and Beck, 2005; Tulin and Cross, 2015). Our data are in agreement that some genes encoding PHCs and HRGPs, including GAS28, were expressed at ZT12 in all three cell strains when cell division takes place, and these genes were further induced in *cht7* following N deprivation (Fig. 2.7A and 2.7B). Together, these findings imply that the expression of some cell wall genes might be suppressed by the same CHT7-mediated regulatory mechanism as other S/M phase genes above, and that CHT7 might play a role opposite to mitotic CDKs which positively regulate the expression of cell wall genes.

In addition to the necessity for cell wall synthesis after cytokinesis, cell wall biosynthesis can also be triggered by damage to the integrity of the cell wall (Cronmiller et al., 2019). Previous research has shown that exposure of cells to gametolysin, a cell wall peptidase normally produced during gametic fusion, induces the expression of cell wall genes necessary for vegetative growth, of which some are also active in the early zygote (Joo et al., 2017). To test if the abnormal induction of cell wall genes in *cht7* is correlated with a misregulated cell wall degradation signal, we examined the expression of cell-wall proteases that might be involved in cell wall remodeling. In the Chlamydomonas genome v5.6, cell wall peptidases such as gametolysin and subtilisin-like proteases are mostly encoded by the MMP (metalloproteases) and SUB (subtilisin) gene families, respectively (Luxmi et al., 2018). We also included genes encoding proteins containing the scavenger receptor cysteine rich (SRCR) domain in our analysis because it has been proposed that SRCR proteins might be involved in the peptidase-mediated cell-wall biosynthesis in Chlamydomonas (Cronmiller et al., 2019; Wheeler et al., 2008). Our results showed that 36 genes encoding MMPs, SUBs, and SRCR proteins were upregulated in *cht7* at more than three timepoints, including ND12, 36, and NR6 (Fig. 2.7A and 2.7C). Notably, the expression patterns of these cell wall peptidases and SRCR protein-encoding genes were very similar to those of cell wall genes (Fig. 2.7B), implying that they belong to the same cell wall remodeling/biosynthetic pathway in Chlamydomonas that appears suppressed by CHT7, especially under long-term ND conditions (Fig. 2.7D). We also found that 30 of these *cht7*-induced genes encoding cell wall remodeling proteins, such as MMP25, MMP31, SUB4, and SRR21, were transcribed at ZT0, 10, or 12 in all three lines (Fig. 2.7C), suggesting that they have an additional function in cell wall reconfiguration during cell division other than gametic fusion. The aberrant transcription profiles in *cht7* are

consistent with previous observations of mis-shapen *cht7* mutant cells under ND and NR conditions (Takeuchi et al., 2020b).

Induction of genes encoding uncharacterized protein kinase families in *cht7*

In addition to the biological processes mentioned above, the GO term GO:0006468 Protein Phosphorylation, was also found to be enriched among proteins encoded by genes that were derepressed in the *cht7* mutant (Fig 3D). In Chlamydomonas, 778 genes have been annotated to encode proteins with this GO term, and 162 of them were upregulated in *cht7* at more than one timepoint. To narrow the scope of our analysis to the most unexplored aspects, we focused on 46 previously uncharacterized, kinase-encoding genes that were induced in *cht7* at more than five timepoints. In order to categorize these kinases and infer possible involvement in signaling pathways, we performed a phylogenetic analysis of the kinase domain of 43 kinases of unknown function with 122 known kinases with characterized functions from Chlamydomonas and various plant species, including Mitogen-activated protein kinase (MAPK) cascade kinases (MAPKs, MAP2Ks, and MAP3Ks) (Doczi et al., 2012; Fei et al., 2017; Krysan et al., 2002; Yu et al., 2007), Calcium-dependent protein kinases (CDPKs, CPKs, CIPKs, CCaMKs) (Cheng et al., 2002; Tirichine et al., 2006; Yu et al., 2007), NEVER IN MITOSIS A (NIMA)-related kinases (CNKs, NEKs, FA2) (Bradley and Quarmby, 2005; Bradley et al., 2004; Mahjoub et al., 2004; Vigneault et al., 2007), Aurora-like kinases (ALKs) (Pan et al., 2004), and a few other kinases known to be associated with flagella and basal bodies (GSK3 and CDKL1) (Tam et al., 2007; Wilson and Lefebvre, 2004).

The phylogenetic analysis in Fig. 2.8A indicated that some *cht7*-induced kinases were related to known kinase families such as MAPK, MAP3K, CIPK, and ALK (Fig. 2.8A; highlighted

in orange). For instance, the Cre17.g10600 and Cre02.g095099-encoding kinases were closely related to Chlamydomonas MAPKKK11 and MAPKKK13 (Fig. 2.8A; MAP3K (1)), while the Cre10.g464100-encoding kinase was related to an Arabidopsis-specific MAP3K family (Fig. 2.8A; MAP3K (2)). In WT and *CHT7-HA::cht7* cells, these MAP3K-related kinases-encoding genes were mainly expressed at ZT10 (Cre17.g10600), ZT12 (Cre02.g095099), or both (Cre10.g464100), but were strongly derepressed in *cht7* at other timepoints (Fig. 2.8B). We also identified two kinases encoded by Cre03.g169100 and Cre17.g735550 that were grouped with the ALK family of Chlamydomonas on the phylogenetic tree (Fig. 2.8A; ALK). Genes encoding these ALK-related proteins and ALK3 were expressed at ZT12 in all strains but greatly derepressed in *cht7* when the cells were deprived of N, and should have ceased division, especially at ND36 (Fig. 2.8C). Lastly, we identified a large group of CHT7-regulated kinase-encoding genes (n=31) that did not show obvious similarities to any of the reference sequences except for the kinase domain, and three clusters were formed based on their sequence similarities (Fig. 2.8A; highlighted in blue, green, and red). In WT and *CHT7HA::cht7*, most of the genes in the three clusters were slightly increased in transcript levels at ZT12 when cells are in the S/M phase, suggesting a potential role of these kinases during the cell cycle (Fig. 2.8D-F). The expression profiles of the genes for these kinases indicated that most cluster 1 (n=17) and cluster 3 (n=8) kinase genes were highly expressed in *cht7* at timepoints besides ZT0 and ND6 and peaked at ZT12 and ND36 (Fig. 2.8D and 2.8F). Genes encoding cluster 2 kinases (n=5) were induced in *cht7* at all timepoints and their expression particularly peaked at ND36 (Fig. 2.8E). In summary, the identification of these potential CHT7-regulated kinase families suggests an uncharacterized, kinase-mediated cell cycle related pathway in addition to the known kinase-mediated signaling pathways in Chlamydomonas.

DISCUSSION

Based on our previous analyses, the CHT7 protein in the green alga Chlamydomonas has been proposed to be involved in the regulation/repression of transcriptional programs particularly relevant during G0 and the transition between G0 and G1 states. To further test this hypothesis, we performed a comprehensive RNA-seq experiment using photoautotrophic cell cultures grown under synchronizing conditions in bioreactors with turbidostatic control allowing us to examine transcriptomic changes caused by the absence of CHT7 at key stages of the cell division cycle and ND-induced quiescence (Fig. 2.2A). A key prerequisite was that, as previously observed (Takeuchi et al., 2020a), most *cht7* cells grown under N+ conditions proceed through the cell division cycle on the same time scale as the WT and *CHT7-HA::cht7*, and the overall growth and viability of the cell culture was not noticeably affected by the absence of CHT7 under N+ growth conditions. However, probing the DNA content of individual cells under these conditions, we observed a subgroup of *cht7* cells with ≥2C DNA content across all non-cell division timepoints (Fig. 2.1A-B, and Fig. S2.1). Moreover, light microscopy at ZT6 revealed that compared to controls, the *cht7* mutant had a more diverse cell population even under N+ growth conditions. For instance, type 2 cells, defined as round and non-flagellated cells, were nearly exclusively found in *cht7* (Fig. 2.1C-D). The phenotypes of the type 2 *cht7* cells resemble those of WT during prophase, with their DNA duplicated, chromosomes condensed, and flagella resorbed (Cross and Umen, 2015). While we were unable to determine whether the type 2 cells of *cht7* were the ≥2C DNA cells described above, overall, these observations suggest that the loss of CHT7 in the mutant allows a fraction of *cht7* cells to escape prematurely the G1 phase of the cell division cycle. Therefore, based on DNA content as a criterion, the *cht7* cell cultures appear to be less synchronized compared to the control strains. In fact, genes encoding proteins involved in cell division, especially DNA replication and

mitotic entry, were abnormally and overwhelmingly more highly expressed in *cht7* at different times throughout the cell cycle in all conditions compared to the WT. A plausible interpretation of these results is that CHT7 is generally needed to suppress expression of genes required for the entry into cell division (Fig. 2.4 and 2.5) and, therefore, its absence could affect the level of synchrony of the culture. It remains to be determined whether spindle formation is abnormally activated under N+ growth of *cht7* cells as well.

Although abnormal cell phenotypes of at least a subpopulation of *cht7* cells were observed under N+ growth conditions, the altered transcriptional profiles raise the question of why *cht7* cells do not exhibit more severe cell cycle defects under normal N+ growth as they do when they are in ND and NR conditions. One plausible reason is that the phenotypes observed for ND and NR *cht7* cells are the result of cumulative effects when cell growth slows under ND and thus are not seen in the rapidly dividing N+ cell cultures. The loss of CHT7, although resulting in a widely misregulated transcriptome in *cht7* even under N+ conditions, did not have an immediate impact on cell cycle progression as observed for known cell cycle disrupted mutants such as *cdkb1* (impaired spindle) and *cdka1* (delayed S/M initiation). This could be because changes in transcript levels in *cht7* are under a threshold that does not immediately affect actual protein abundance or activity, or due to the fact that only a subset of cell cycle genes was induced (Fig. 2.6). For example, APC is a conserved cell cycle regulatory complex in eukaryotes and is required for mitotic exit by promoting chromosome segregation and mediating the degradation of mitotic cyclins (Cross and Umen, 2015). The Chlamydomonas mutant lacking APC3, *cdc27-6*, arrests at metaphase, indicating the essential role of APC in the completion of mitosis (Atkins and Cross, 2018). The genes encoding the APC components were expressed normally in *cht7* at all ZT timepoints (Fig. 2.6), although some were misregulated during ND. Accordingly, although DNA replication and

some mitotic events are abnormally activated in *cht7* cells during G0 and G1, critical genes for cell division in *cht7* may allow the S/M phase to proceed when the cells are not under metabolic stress. However, during an extended period following ND, the imbalanced expression of many cell cycle genes in *cht7* and the abnormally activated S/M phase processes may eventually lead to irreversible metabolic outcomes, abnormal number of nuclei and arrested cell division (Takeuchi et al., 2020b).

Another possible reason for the relatively mild phenotype of N+ *cht7* cells compared to ND and NR cells is that CHT7 might govern additional pathways during quiescence. For instance, the apparent increase in cell wall synthesis or cell wall remodeling in *cht7* at ND time points might be responsible for the cellular phenotypes exhibited in N-deprived *cht7* cells (Fig. 2.7). Of note, previous studies using the cell wall-less (cw-) and cell-walled (cw+) *cht7* mutant strains observed different phenotypes: the viability of cw+ *cht7* cells declined to 40-50% following an extended period of ND, while the cw- *cht7* cell did not exhibit a loss of viability (Takeuchi et al., 2020b; Tsai et al., 2014). The basis of the differential viability may be deduced from the data reported here: derepression of the cell wall pathways may contribute to the increased mortality of the cw+ *cht7* mutant during quiescence since the Chlamydomonas cell wall would be a major sink for cellular nitrogen and energy (Lang and Chrispeels, 1976); continuously active cell wall remodeling might be harmful to the cell, especially under N-limited conditions. Based on the insights from our current analysis, a future study analyzing the cell wall structure and remodeling of *cht7* cells should provide answers to the question of differences in apparent viability between the cw+ and cw- *cht7* mutant strains.

In addition to the genes affected by CHT7 that have defined functions, we identified several previously uncharacterized genes that are misregulated in *cht7* and may encode kinases related to

MAPK cascade kinases or aurora-like kinases (Fig. 2.8A). These genes are highly expressed at ZT10-12 (Fig. 2.8B and 2.8C), suggesting they may function in cell division. Three uncharacterized kinase families were identified in our dataset (Fig. 2.8A, 2.8D-F), whose genes were strongly derepressed in *cht7* compared to the two controls. These should provide promising candidates for further investigations of cell division regulation in Chlamydomonas.

Finally, this study aimed to provide insight into the possible role of CHT7 in cell cycle regulatory mechanisms of Chlamydomonas. We showed that CHT7 is required for the repression of specific groups of cell cycle genes not only during quiescence, but also in the normal cell division cycle (Fig. 2.3). As mentioned previously, CHT7 impacts the expression of S-phase and mitosis-related genes, however, the mechanism by which CHT7 represses the expression of these genes remains to be determined. As reported previously, CHT7-based ChIP-Seq analysis did not detect a direct interaction between CHT7 and chromosomal DNA (Takeuchi et al., 2020b). Subsequently, it has been shown that the presumed DNA binding domain (CHC domain) of the CHT7 protein was not required for its function during N deprivation and following N resupply (Takeuchi et al., 2020b). Together, these results strongly suggested that CHT7 does not act as a canonical transcriptional repressor by binding directly to DNA through its CHC domain. However, one could postulate that CHT7 interacts with other transcriptional repressors that directly bind to the respective target genes. The identification of the primary targets of CHT7 in complex with other protein factors will be needed to clarify this point. The HA-CHT7 protein having the deletion of the disordered region would serve as a negative control for such protein-protein interaction assay since that directed deletion creates a null mutation (Takeuchi et al., 2020b). This next step in the analysis of CHT7 function would also help resolve uncertainties affecting the interpretation of the current data such as the possibility that the induced expression of cell cycle genes in *cht7*

was a compensatory effect due to other factors such as compromised central metabolism or partial loss of viability, or that other genes disrupted in the *cht7* mutant contribute to the observed phenotypes.

Aside from these uncertainties, our current data support a hypothesis that assigns CHT7 a role in the regulation of the cell cycle opposite to mitotic CDKs in Chlamydomonas. It has been shown that CDKA1 participates in the initiation of S-phase and cell division, while CYCB1-CDKB1 plays a more important role in promoting the entry into mitosis (Atkins and Cross, 2018). Both proteins are in the nucleus and function by protein phosphorylation affecting transcriptional activation of the respective genes. In addition, CDKA1 also induces the expression of *CYCB1* and *CDKB1* (Atkins and Cross, 2018; Tulin and Cross, 2015). In *cht7*, genes encoding CDKB1 and CYCB1 were derepressed in a similar pattern to other misregulated cell cycle genes, while the transcription of the CDKA1-encoding gene was unaffected (Fig. 2.6). This result raised the possibility that a CHT7 containing complex may affect the expression of similar genes as CDKA1, but negatively regulates their expression to inhibit cell division when not desirable. A double mutant analysis of *cht7* and *cdka1* asking whether *cht7* suppresses effects of *cdka1* would be helpful to clarify this point. The current work should inspire future efforts towards the identification of CHT7-associated transcription factors and functional characterization of previously unreported protein kinases to gain a more complete understanding of cell cycle regulation in Chlamydomonas.

MATERIALS AND METHODS

Cell strains and growth conditions

Initial characterization of the three mating-type minus, cell-walled *Chlamydomonas reinhardtii* strains used in this study, 6145c (wild type), *cht7*, and the *cht7* complementation line *CHT7*-

HA::cht7, was reported previously (Takeuchi et al., 2020b). Synchronization of Chlamydomonas cells in HS medium with an environmental photobioreactor system ePBR using turbidostatic conditions (Lucker et al., 2014) has been described in detail by Takeuchi et al. (2020a. Three cell-walled, mating-type minus strains were used in this study: wild type (WT, 6145c), the *cht7* mutant, and the *cht7* complemented line *CHT7-HA::cht7*. The complemented line was included here as an additional control to detect any possible effects caused by the deletion of genes other than CHT7 and other possible background mutations in the *cht7* mutant (Takeuchi et al., 2020b). For synchronization of Chlamydomonas cells in environmental photobioreactors (ePBRs) (Lucker et al., 2014; Takeuchi et al., 2020a), cells were grown autotrophically in 330 mL of nitrogen containing high salt (HS+N) medium (Sueoka, 1960) in ePBR under 12 h light: 12 h dark cycles with a light intensity of 2,000 μmol $m^{-2}s^{-1}$, and 5% CO_2 was pumped into the cell culture regularly (30 seconds per 15 minutes). To maintain cell growth in mid-log phase, OD_{940} of the cell culture was monitored and diluted automatically by the ePBR using its turbidostat function. The target OD_{940} was determined according to the chlorophyll concentration (μg/mL) at ZT6, and all cultures were normalized to around 3 μg/mL of total chlorophyll. The cells culture was grown under the condition above for 7-8 days to be synchronized prior to sample harvesting.

For N deprivation of the synchronized ePBR cell cultures, at ZT23 cells of each culture were pelleted at 3,000 g for 5 minutes and rinsed twice with 50 mL of N-free HS (HS-N) medium before being resuspended with 330 mL of HS-N medium. The N-deprived cells were then transferred to a clean ePBR vessel and grown under 2,000 μmol $m^{-2}s^{-1}$ of constant light with other growth conditions the same as the N+ cultures. For N-resupply of the ND culture, aliquots of a sterile NH_4Cl solution were added into N-deprived ePBR cultures at ND36 to bring the final concentration of NH_4Cl to be the same as the HS+N medium.

Flow cytometry analysis

Sample pretreatment for flow cytometry analysis was conducted according to a previous study (Fang et al., 2006). 5 mL of cells harvested from ePBR culture were pelleted and resuspended in 5 mL of Ethanol/Acetic acid (3: 1) solution for fixation for 2 hours at room temperature. After fixation, cells were rinsed twice and resuspended with 1 mL of FACS buffer containing 0.2M Tris-HCl (PH 7.5) and 20mM EDTA and stored at 4°C. Prior to flow cytometry analysis, the cell density of each sample was measured with a Z2™ COULTER COUNTER® Analyzer (Cat. 6605700, Beckman Coulter), and all samples were normalized to 1×10^6 cells/mL in a total 900 µL of FACS buffer, mixed with 100 µL of FACS buffer containing 1 mg/mL of RNase A, and incubated at 37°C for 2 hours. For DNA staining, the RNase A-treated samples were rinsed once and resuspend in 1 mL of PBS buffer. 495 µL of sample were mixed with 5 µL of 100-fold diluted SYTOX™ Green (Cat. S7020, Thermo Fisher Scientific) for 5 minutes. Flow cytometry was performed with the LSR II Flow Cytometer System (BD Biosciences) at MSU Flow Cytometry Core (https://facs.iq.msu.edu/), and data were analyzed with FACSDiva™ Software (BD Biosciences).

RNA sequencing

Samples were harvested from ePBR cell cultures at various timepoints. Total RNA was extracted with Rneasy® Plant Mini Kit (cat. 74903, Qiagen). The RNA concentration was measured with Qubit® RNA BR Assay and Fluorometer (Cat. Q10210 and Cat. Q32866, Thermo Fisher Scientific). The quality of RNA was determined with Agilent 2100 Bioanalyzer and Agilent RNA 6000 Pico Assay (Cat. G2939BA and Cat. 5067-1514, Agilent Technologies) to ensure an RNA Integrity Number (RIN) value ≥8 before submitting an average of 4.36 µg of total RNA (220 ng/µl) per sample for RNA sequencing. Lastly, RNA sequencing was done in the Genomics Core at MSU

RTSF (https://rtsf.natsci.msu.edu/genomics/). The cDNA libraries were prepared using TruSeq Stranded mRNA Library Preparation Kit (Illumina) and quantified using the Kapa Biosystems Illumina Library Quantification qPCR kit (Illumina). The pools of ZT sample and ND/NR sample libraries were loaded onto a total of four lanes of Illumina HiSeq 4000 flow cells and sequenced in a 1x50 bp single-end mode using HiSeq 4000 SBS reagents. For each sample, around 25 million reads were generated, processed, and then mapped to Creinhardtii_281_v5.6.transcript downloaded from Phytozome v12.1 (http://phytozome.jgi.doe.gov) using Salmon 0.11.3 (Patro et al., 2017) and R packages tximport 1.16.1 (Soneson et al., 2015) and DESeq2 1.28.1 (Love et al., 2014) to obtain the normalized read count of individual transcripts.

Data processing and differential expression analysis

The raw sequence reads were mapped to Creinhardtii_281_v5.6 and calculated for transcript abundance with Salmon 0.11.3 (Patro et al., 2017). The quantified read counts were normalized through division by the size factor of each sample with the DESeq2 1.28.1 package in R v4.0.0 (Love et al., 2014). Principal Component Analysis (PCA) of all 81 samples was done using the same package with the variance stabilizing transformation (VST) function. The average read count of individual transcripts was calculated from 3 biological replicates of each group (e.g. WT_ZT6). Differential expression analysis was performed to obtain the $\log_2$ fold change of replicative average read counts of a transcript of any two strains at the same timepoint (e.g. Cre08.g372550.t1.1, *cht7* vs. wild type at ND6). To estimate the statistical significance of the fold change, adjusted *p*-value (padj) was calculated from the original *p*-value using the Benjamini–Hochberg rule to correct for the false discovery rate (false positive). Differentially expressed genes (DEGs) were identified using padj < 0.05 and abs (log2 fold change) > 1. A *cht7*-specific DEG is

defined as a gene being differentially expressed in *cht7* compared to both WT and *CHT7-HA::cht7*, but not being differentially expressed between WT and *CHT7-HA::cht7*.

Hierarchical clustering

Among the total 19,526 transcripts, we selected 14,456 transcripts with an average read count > 20 that were differentially expressed (abs (log2 fold change) > 1, padj < 0.05) between any timepoint and the mean of all timepoints. Read counts of these DEGs were then scaled to Z-score by the equation $x\text{-}\bar{x}/\sigma$, where x is the read count of a timepoint and $\bar{x}$ is the mean of all timepoints, σ is the standard deviation of the mean. Hierarchical clustering of the scaled data matrix was then performed in R using the Pearson's correlation option to measure the similarity in Z-score across all timepoints between genes, and 14 gene clusters were generated by cutting the resulting hierarchical tree at height = 1.7. As a result, genes of each cluster have a unique expression profile in response to various genotypes and conditions.

GO enrichment analysis

Two references of GO annotation downloaded from EnsemblPlants (dataset: creinhardtii_eg_gene) and Phytozome v12.1(dataset: Creinhardtii_281_v5.6) were combined into one GO annotation table and further analyzed in R. Among the total of 17,743 genes in the Chlamydomonas genome v5.6, 11,159 genes were annotated with at least one GO term. The input gene set for GO enrichment analysis was composed of the 14 gene clusters (Fig 2C) or the *cht7*-specific upregulated DEGs (Fig 3C). The analysis was performed using the R package topGO 2.40.0 (Alexa and Rahnenfuhrer, 2020; Alexa et al., 2006) with the *nodeSize = 5* option to construct the GO tree prior to analysis. The significance of enriched GO terms was determined by Fisher's exact test

combining with the *weight01* algorithm to determine the most representative GO term by comparing the enrichment score of a GO term with its subsequent GO term. The threshold for a significantly enriched GO term is p-value < 0.05.

Phylogenetic analysis

Peptide sequences of the Chlamydomonas kinases including both the 46 *cht7*-induced kinases and reference kinases were obtained from Phytozome v12.1 (dataset: Creinhardtii_281) using R package biomaRt 2.44.0 (Durinck et al., 2009), and sequences of Arabidopsis kinases were downloaded from Phytozome v12.1 using PhytoMine web tools. Sequences of CcaMKs were obtained directly from Tirichine et al. (2006. Three *cht7*-induced kinases encoded by Cre01.g055457.t2.1, Cre12.g526150.t1.2, and Cre12.g526450.t1.2 were removed from the analysis to avoid duplicated sequences. Sequences of kinase domain were retrieved using an R script searching for IPR000719 Protein kinase domain defined by InterPro (https://www.ebi.ac.uk/interpro/). For the kinase encoded by Cre09.g404550.t1.2, only the IPR004166 MHCK/EF2 kinase domain was found and was used in this study. For the kinase encoded by Cre02.g108750.t1.2, two kinase domains were found, and they were both included in the analysis (labeled as Cre02.g108750a and Cre02.g108750b). Multiple sequence alignment of the total 166 kinase domains (122 reference sequences and 44 sequences of interest) above was performed using ClustalOmega in R package msa 1.19.0 (Bodenhofer et al., 2015). Maximum likelihood (ML) phylogeny was estimated using the Whelan and Goldman (WAG) substitution model implemented in PhyML 3.3 software (Guindon et al., 2010; Whelan and Goldman, 2001), and branch support was measured by a parametric, Chi2-based approximate Likelihood-Ratio

(aLRT) test (Anisimova and Gascuel, 2006). The resulting unrooted ML tree was furthered analyzed and rooted with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

Accession number

The transcriptomic data including the raw and processed sequence reads are deposited to GEO (Gene Expression Omnibus; https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE167405.

ACKNOWLEDGMENTS

This work was primarily supported by grants from the National Science Foundation, MCB-1515169 to C.B. and MCB 1515220 to J.U. In addition, C.B. was supported by MSU AgBioResearch and a grant from the Chemical Sciences, Geoscience and Biosciences Division, Office of Basic Energy Sciences, Office of Science, U.S. Department of Energy (award number DE-FG02-91ER20021). The contribution by J.U. was also supported by NIH grant R01 GM126557. We thank Dr. Daniel Vocelle from MSU Flow Cytometry Core Facility (https://facs.iq.msu.edu/) for providing technical support.

APPENDIX

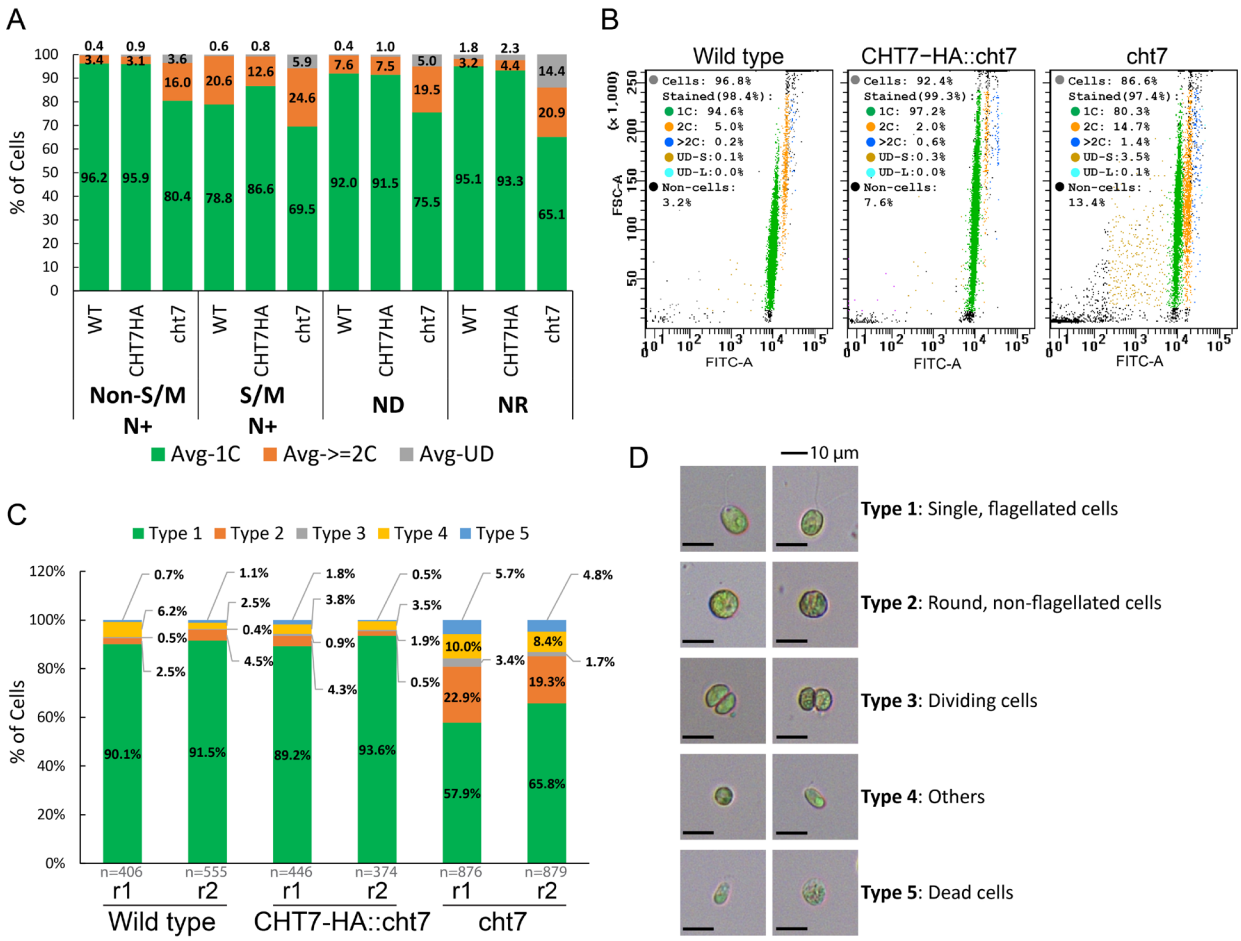
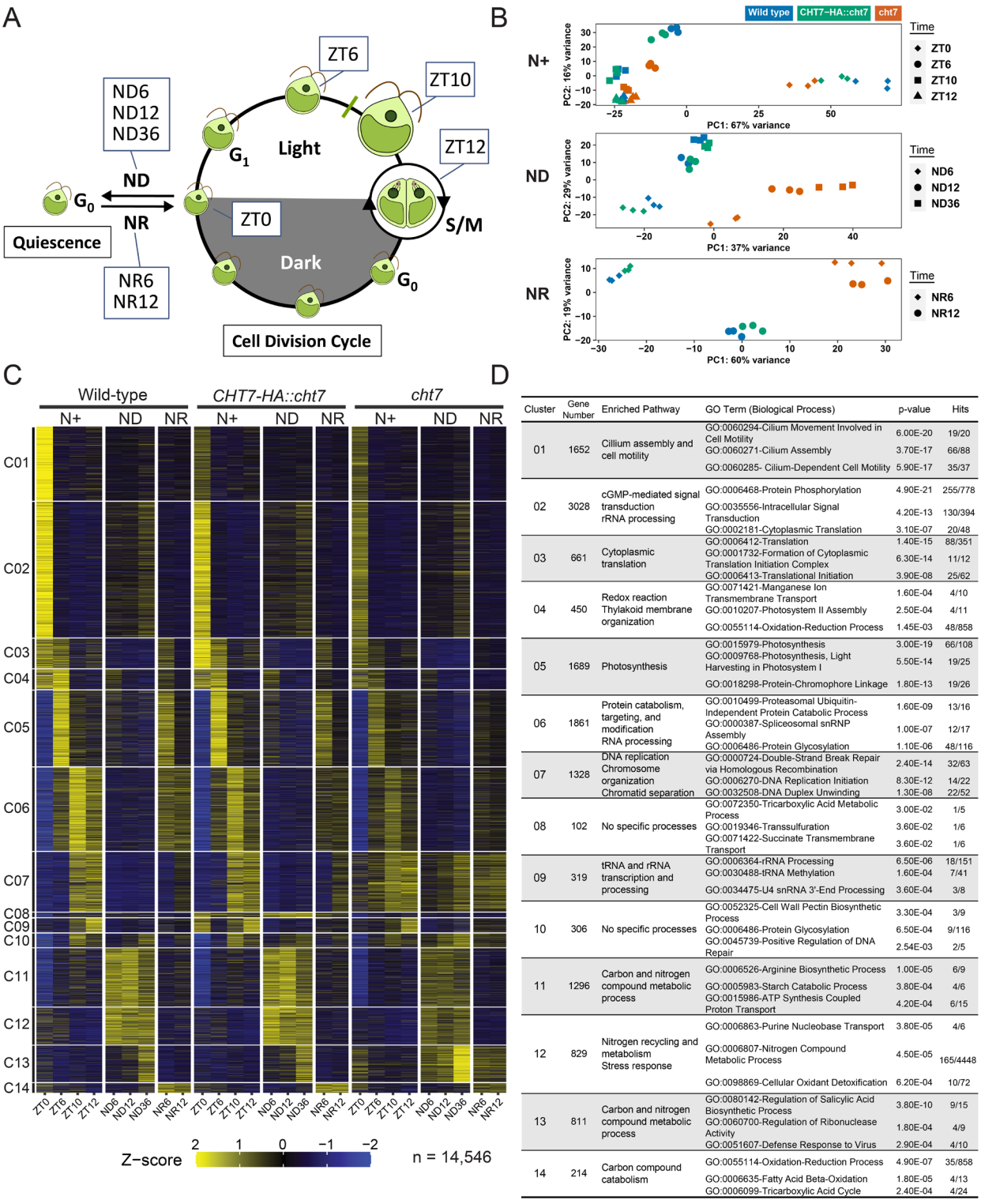


Figure 2.1. Assessment of cell cycle progression under synchronizing conditions.

(A) and ***(B)*** Flow cytometry analysis of DNA-stained cell cultures grown under N+, ND, and NR conditions. Samples of cells grown under synchronizing conditions were fixed and stained with SYTOX Green prior to flow cytometry analysis. Summary of DNA content ***(A)*** of cells at non-S/M (ZT0-10, ZT24) and S/M phases (ZT12-16) of N+ condition as well as cells under ND (ND0, 6, 12, and 36) and NR (NR6 and 12) conditions. *Y*-axis: average population (%) of cells carrying single (green), ≥2 (orange), or undetermined (grey) copies of chromosomes. n=2 for each sample. Example plot ***(B)*** showing particle size (FSC-A; Forward scatter) vs. DNA content (FITC-A) for samples at ZT6. Each dot represents non-cell particle (black) or a cell carrying a single (green), 2

(orange), >2 (blue), or undetermined (brown and cyan) copies of the genome. One of the two replicates (n=2) for each sample is presented. See also Figure S2.1. ***(C)*** and ***(D)*** Cell Morphology of wild type, *CHT7-HA::cht7*, and *cht7* cells at ZT6. Two biological replicates (r1 and r2) for each strain were observed under the microscope and summarized in ***(C)***. n: numbers of cells observed. *Y*-axis: population (%) of cells of type 1 (green), 2 (orange), 3 (grey), 4 (yellow), and 5 (blue). The representative graphs and definition of each cell type are shown in ***(D)***.



Cluster	Gene Number	Enriched Pathway	GO Term (Biological Process)	p-value	Hits
01	1652	Cillium assembly and cell motility	GO:0060294-Cilium Movement Involved in Cell Motility	6.00E-20	19/20
			GO:0060271-Cilium Assembly	3.70E-17	66/88
			GO:0060285- Cilium-Dependent Cell Motility	5.90E-17	35/37
02	3028	cGMP-mediated signal transduction rRNA processing	GO:0006468-Protein Phosphorylation	4.90E-21	255/778
			GO:0035556-Intracellular Signal Transduction	4.20E-13	130/394
			GO:0002181-Cytoplasmic Translation	3.10E-07	20/48
03	661	Cytoplasmic translation	GO:0006412-Translation	1.40E-15	88/351
			GO:0001732-Formation of Cytoplasmic Translation Initiation Complex	6.30E-14	11/12
			GO:0006413-Translational Initiation	3.90E-08	25/62
04	450	Redox reaction Thylakoid membrane organization	GO:0071421-Manganese Ion Transmembrane Transport	1.60E-04	4/10
			GO:0010207-Photosystem II Assembly	2.50E-04	4/11
			GO:0055114-Oxidation-Reduction Process	1.45E-03	48/858
05	1689	Photosynthesis	GO:0015979-Photosynthesis	3.00E-19	66/108
			GO:0009768-Photosynthesis, Light Harvesting in Photosystem I	5.50E-14	19/25
			GO:0018298-Protein-Chromophore Linkage	1.80E-13	19/26
06	1861	Protein catabolism, targeting, and modification RNA processing	GO:0010499-Proteasomal Ubiquitin-Independent Protein Catabolic Process	1.60E-09	13/16
			GO:0000387-Spliceosomal snRNP Assembly	1.00E-07	12/17
			GO:0006486-Protein Glycosylation	1.10E-06	48/116
07	1328	DNA replication Chromosome organization Chromatid separation	GO:0000724-Double-Strand Break Repair via Homologous Recombination	2.40E-14	32/63
			GO:0006270-DNA Replication Initiation	8.30E-12	14/22
			GO:0032508-DNA Duplex Unwinding	1.30E-08	22/52
08	102	No specific processes	GO:0072350-Tricarboxylic Acid Metabolic Process	3.00E-02	1/5
			GO:0019346-Transsulfuration	3.60E-02	1/6
			GO:0071422-Succinate Transmembrane Transport	3.60E-02	1/6
09	319	tRNA and rRNA transcription and processing	GO:0006364-rRNA Processing	6.50E-06	18/151
			GO:0030488-tRNA Methylation	1.60E-04	7/41
			GO:0034475-U4 snRNA 3'-End Processing	3.60E-04	3/8
10	306	No specific processes	GO:0052325-Cell Wall Pectin Biosynthetic Process	3.30E-04	3/9
			GO:0006486-Protein Glycosylation	6.50E-04	9/116
			GO:0045739-Positive Regulation of DNA Repair	2.54E-03	2/5
11	1296	Carbon and nitrogen compound metabolic process	GO:0006526-Arginine Biosynthetic Process	1.00E-05	6/9
			GO:0005983-Starch Catabolic Process	3.80E-04	4/6
			GO:0015986-ATP Synthesis Coupled Proton Transport	4.20E-04	6/15
12	829	Nitrogen recycling and metabolism Stress response	GO:0006863-Purine Nucleobase Transport	3.80E-05	4/6
			GO:0006807-Nitrogen Compound Metabolic Process	4.50E-05	165/4448
			GO:0098869-Cellular Oxidant Detoxification	6.20E-04	10/72
13	811	Carbon and nitrogen compound metabolic process	GO:0080142-Regulation of Salicylic Acid Biosynthetic Process	3.80E-10	9/15
			GO:0060700-Regulation of Ribonuclease Activity	1.80E-04	4/9
			GO:0051607-Defense Response to Virus	2.90E-04	4/10
14	214	Carbon compound catabolism	GO:0055114-Oxidation-Reduction Process	4.90E-07	35/858
			GO:0006635-Fatty Acid Beta-Oxidation	1.80E-05	4/13
			GO:0006099-Tricarboxylic Acid Cycle	2.40E-04	4/24

Figure 2.2. Transcriptome analysis of Chlamydomonas cell cultures under synchronizing conditions

Figure 2.2. (cont'd)

(A) Diagram of Chlamydomonas cell cycle and experimental design for RNA-sequencing analysis. Wild type, *CHT7-HA::cht7*, and *cht7* cells were grown photoautotrophically in an environmental photobioreactor (ePBR) under turbidostatic control. ZT(n) samples represent cells at different stages of the cell division cycle under normal N+ conditions, where n = hours after the onset of light. Quiescence was triggered by nitrogen deprivation (ND) in continuous light. ND(n) samples were harvested, where n = hours after N deprivation. The exit from quiescence was induced by N resupply (NR). NR(n) samples were harvested after n hours of N refeeding. Three independent biological samples (*n*=3) were analyzed. ***(B)*** Principal component analysis (PCA) of transcriptomes of all 81 samples. Samples are labeled with different colors and shapes as shown on the top and right of the graph to indicate combinations of different genotypes and timepoints, respectively. ***(C)*** Z-score heatmap and ***(D)*** GO enrichment analysis of 14 gene clusters across all conditions. 14,456 transcripts with an average normalized read count > 20 (n=3) differentially expressed in any sample and timepoint were grouped into 14 gene clusters by hierarchical clustering based on their Z-scores. Each cluster represents a unique expression pattern across all conditions. Hits (x/y): number of the input genes (x) over total genes (y) within a GO term. See also Figure S2.1.

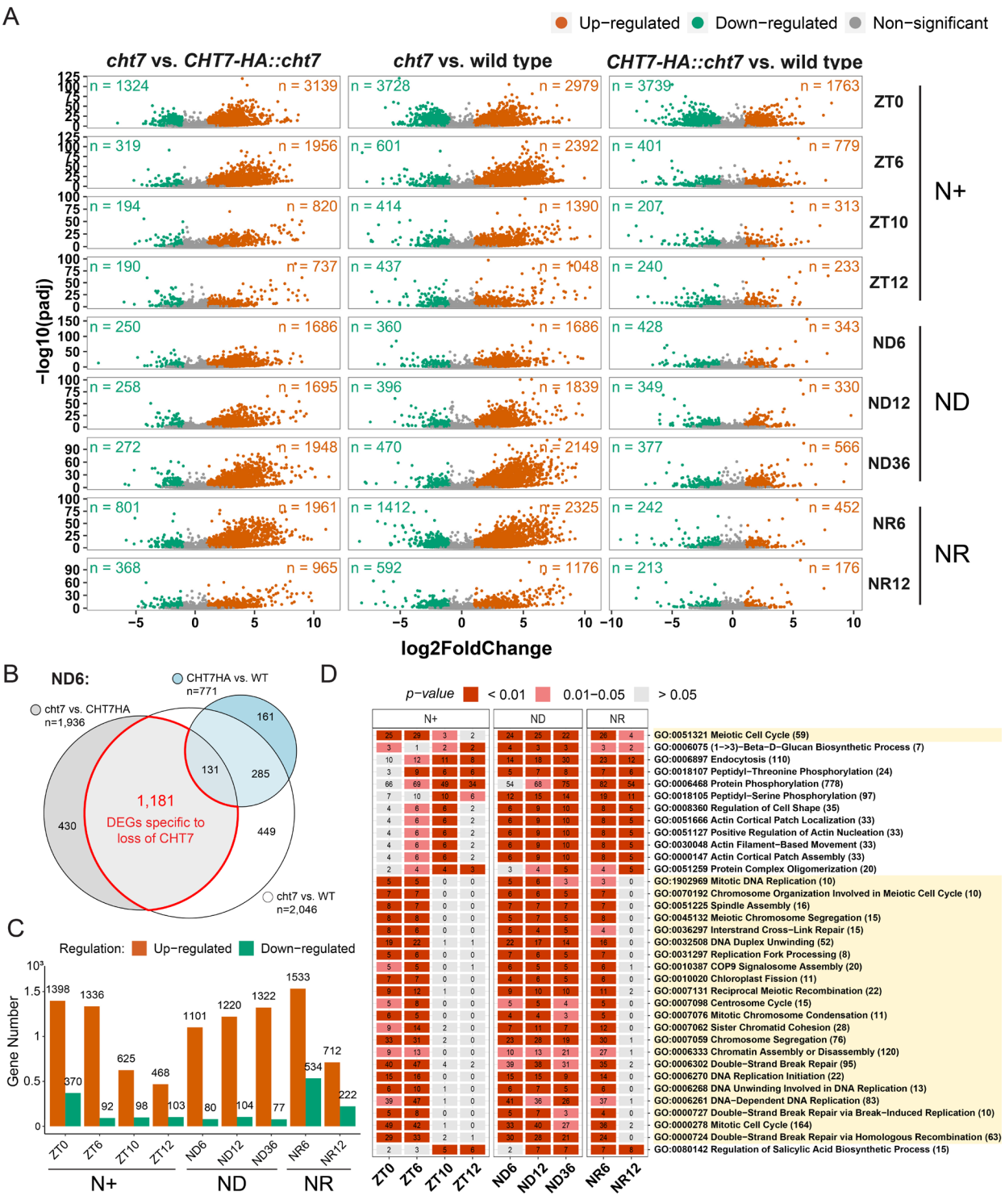


Figure 2.3. Differential expression analysis of all three strains.

Figure 2.3 (cont'd)

(A) Volcano plots showing the upregulated (orange; $\log_2$ fold change > 1, *padj* < 0.05), down-regulated (green; $\log_2$ fold change < 1, *padj* < 0.05), and non (gray) DEGs in three groups of any two genotypes (e.g. *cht7* vs. wild type) across all timepoints. *X*-axis: $\log_2$ fold change. *Y*-axis: -$\log_{10}$ adjusted *P*-value (*padj*). n= number of DEGs. ***(B)*** Venn diagram example showing the selection of *cht7*-specific DEGs from the ND6 samples. DEGs between the WT and *CHT7HA::cht7* were excluded from DEGs between *cht7* and WT and *CHT7HA::cht7*, respectively . ***(C)*** Numbers of up- (orange) and down- (green) regulated *cht7*-DEGs across all timepoints. ***(D)*** GO enrichment analysis of upregulated *cht7*-DEGs across all timepoints. 23 GO terms of biological processes related to the S/M phase are highlighted in light yellow. Number of upregulated DEGs are labeled in the boxes. Number of total annotated genes are indicated after GO terms. See also Figure S2.2.

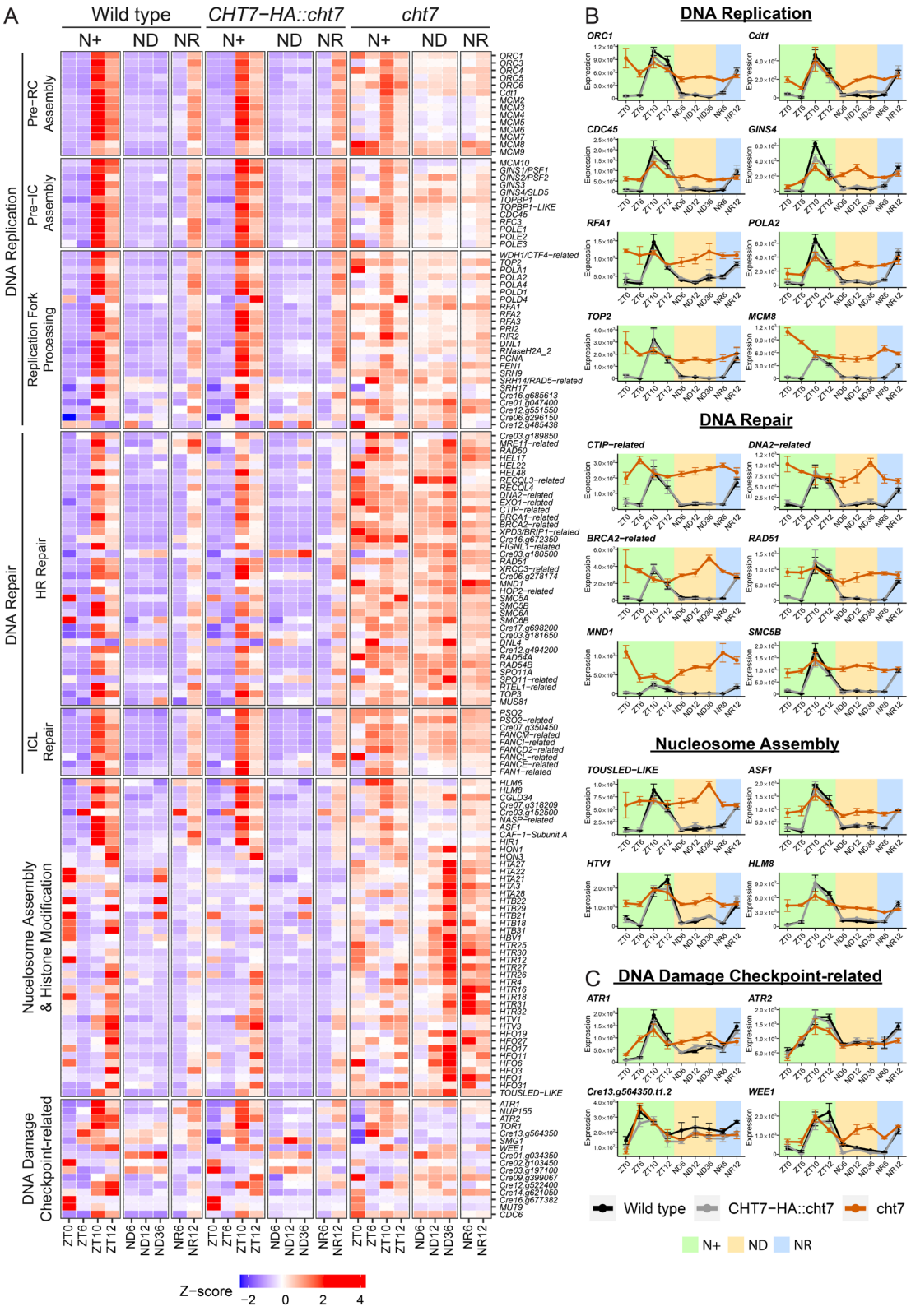


Figure 2.4. Derepressed DNA replication in the cht7 mutant.

Figure 2.4 (cont'd)

(A) Z-score heatmap of S-phase genes across all conditions. Genes that were differentially expressed ($\log_2$ fold change > 1, *padj* < 0.05) in *cht7* at ≥ 1 timepoints are presented. (pre-RC) Pre-replicative complex; (pre-IC) Pre-initiation complex; (HR) Homologous recombination repair; (ICL) Interstrand cross-link. ***(B)*** and ***(C)*** Expression of representative S-phase genes ***(B)*** and DNA damage checkpoint-related genes ***(C)*** in *cht7* (orange), WT (black), and *CHT7-HA::cht7* (gray) across all timepoints. *Y*-axis: normalized read count. (ORC1) Origin Recognition Complex 1; (Cdt1) Cdc10-dependent transcript 1; (CDC45) Cell Division Cycle 45; (RFA1) Replication Factor A1; (POLA2) DNA Polymerase α2; (TOP2) Topoisomerase II; (MCM8) Minichromosome Maintenance 8; (CTIP) CtBP (C-terminal Binding Protein) Interacting Protein; (BRCA2) Breast Cancer gene 2. (MND1) Meiotic Nuclear Divisions 1; (SMC5B) Structural Maintenance of Chromosomes 5; (ASF1) Anti-Silencing Factor 1; (HTV1) Histone H3 variant 1; (HLM8) Histone-lysine N-methyltransferase 8; (ATR) Ataxia Telangiectasia Mutated (ATM) and Rad3-related.

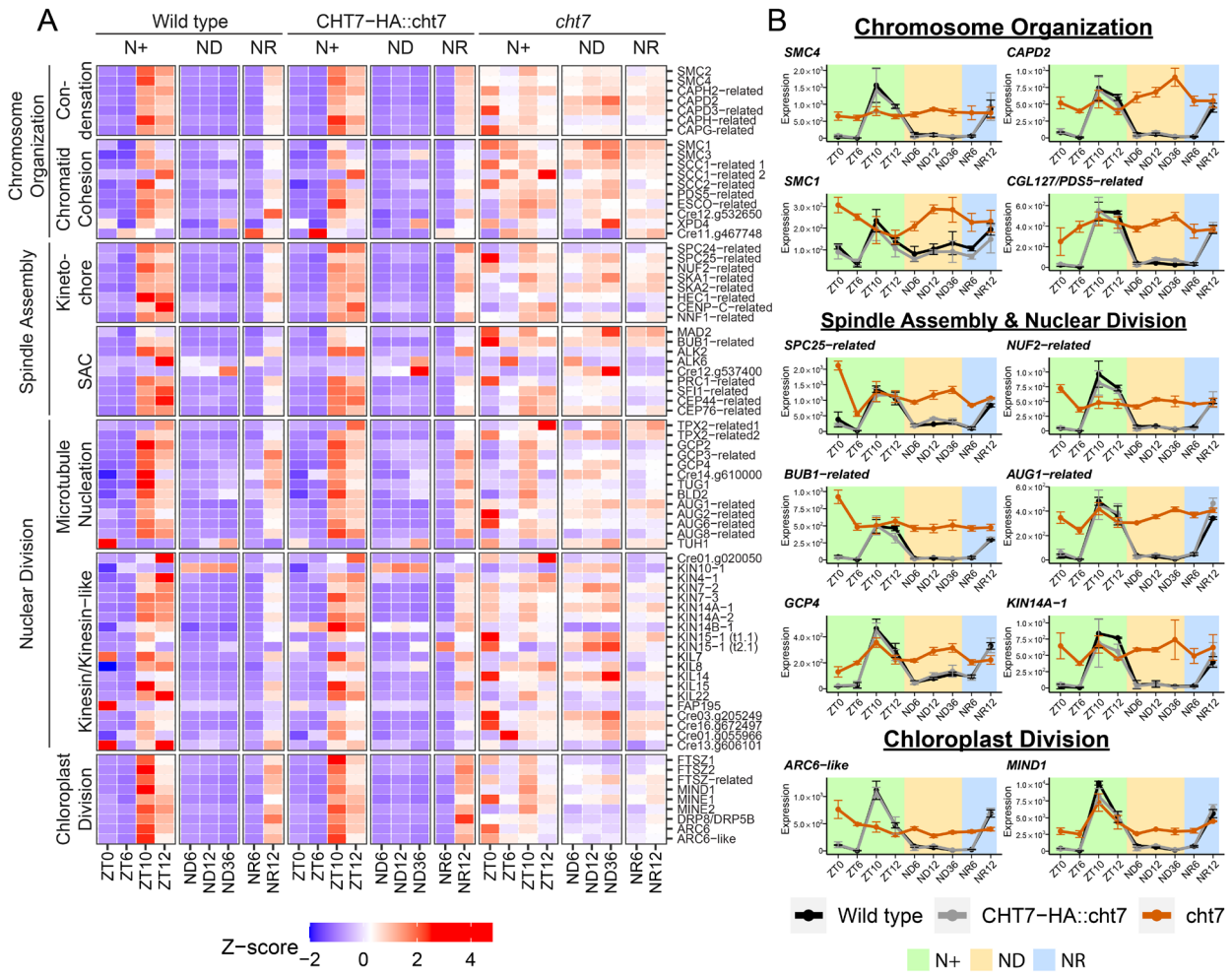


Figure 2.5. Derepressed mitotic genes in the *cht7* mutant.

(A) Z-score heatmap of mitotic genes. Genes that were differentially expressed (log_2 fold change > 1, *padj* < 0.05) in *cht7* at ≥ 1 timepoints are presented. (SAC) Spindle assembly checkpoint ***(B)*** Expression of representative mitotic genes in *cht7* (orange), WT (black), and *CHT7-HA::cht7* (gray) across all timepoints. *Y*-axis: normalized read count. (SMC4) Structural Maintenance of Chromosomes 4; (CAPD2) Chromosome-Associated Protein D2; (SPC25) Spindle Pole Component 25; (NUF2) Nuclear Filament-containing protein 2; (AUG1) AUGMIN subunit 1; (BUB1) Budding Uninhibited by Benzimidazole 1; (GCP4) γ-tubulin Complex Protein 4; (KIN14A-1) Kinesin-like Protein 14A-1; (ARC6-like) Accumulation and Replication of Chloroplasts 6-like; (MIND1) Minicell mutant D1.

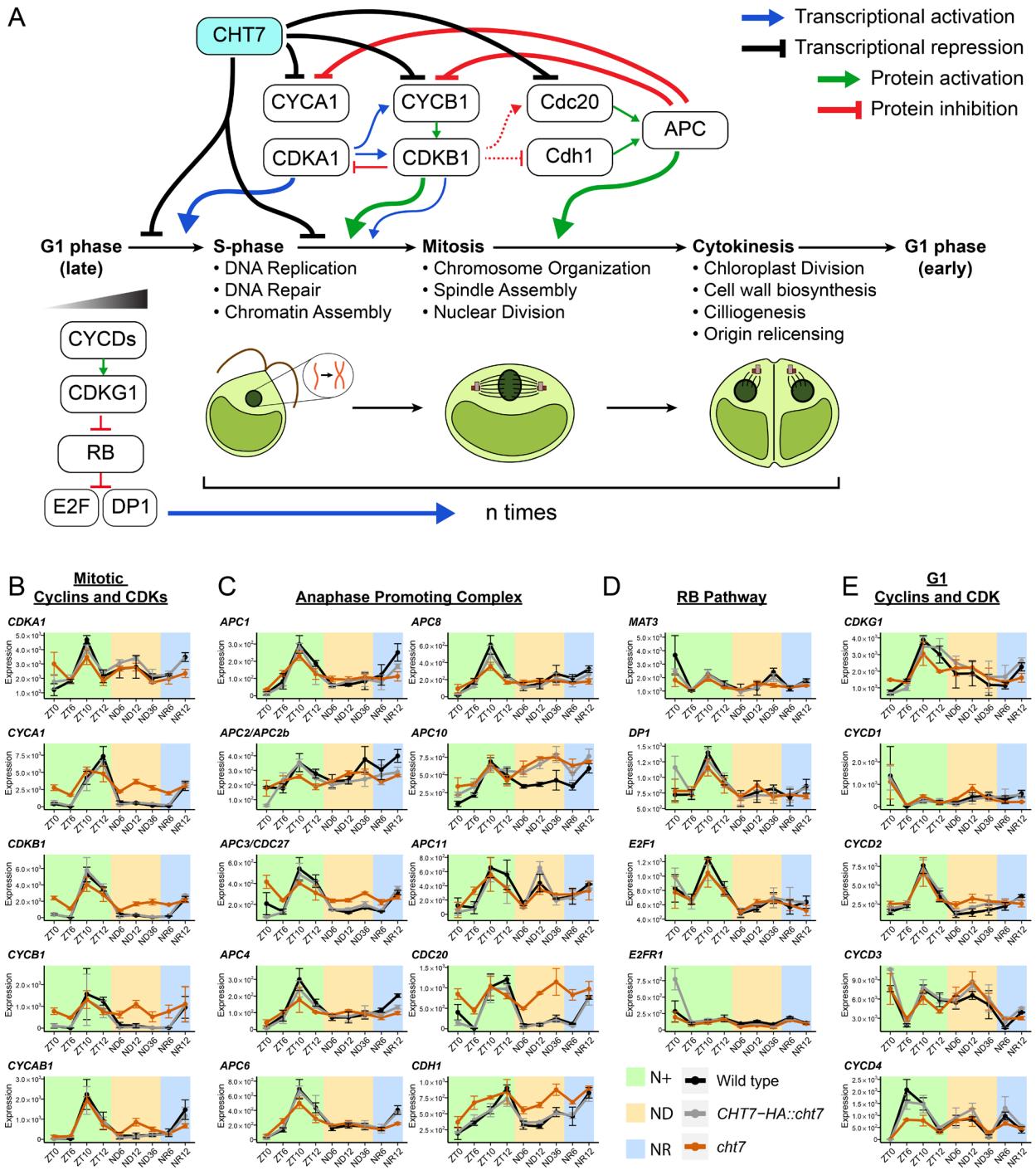


Figure 2.6. Impact of CHT7 on cell cycle regulation of Chlamydomonas.

(A) Regulatory network of the cell cycle of Chlamydomonas and proposed model of CHT7-mediated cell cycle regulation. Refer to text for details. ***(B-E)*** Expression of the genes encoding

Figure 2.6 (cont'd)

cell cycle regulators in *cht7* (orange), WT (black), and *CHT7-HA::cht7* (gray) across all timepoints. *Y*-axis: normalized read count. (RB) Retinoblastoma protein; (E2F) E2 factor; (DP1) Dimerization Partner 1; (CYC) cyclin; (CDK) Cyclin-Dependent Kinase; (APC) Anaphase Promoting Complex; (CDC20) Cell Division Cycle protein 20; (CDH1) Cdc20 Homologue-1.

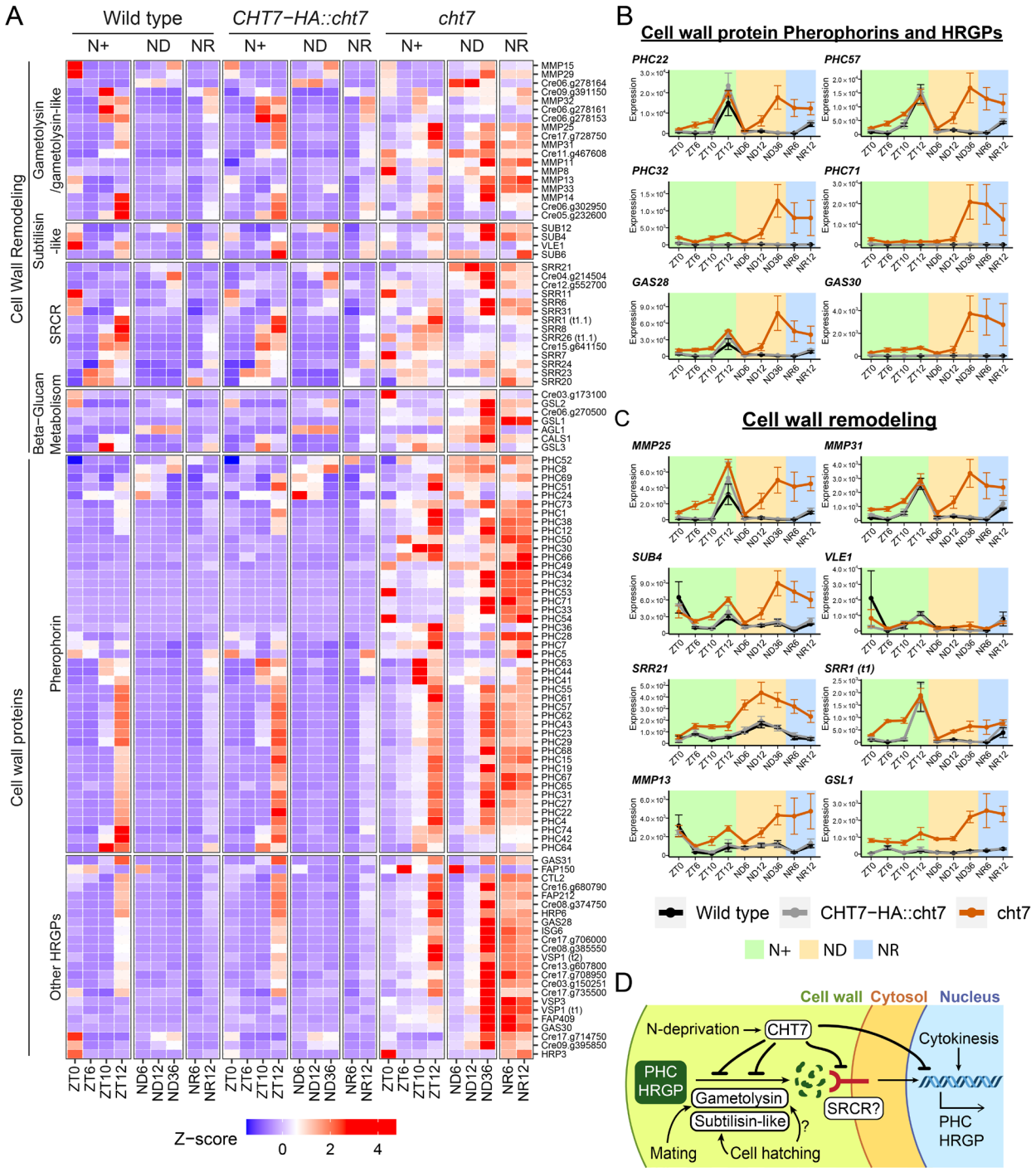


Figure 2.7. Misregulated cell wall-associated genes in the *cht7* mutant.

(A) Z-score heatmap of cell wall remodeling and biosynthetic genes. Genes that were differentially expressed ($\log_2$ fold change > 1, *padj* < 0.05) in *cht7* at ≥ 3 timepoints are presented. (SRCR) Scavenger receptor cysteine-rich. ***(B-C)*** Expression of representative genes in *cht7* (orange), WT

Figure 2.7 (cont'd)

(black), and *CHT7-HA::cht7* (gray) across all timepoints. *Y*-axis: normalized read count. (PHC) Pherophorin; (MMP) Matrix Metalloprotease; (SUB) Subtilisin-like protease; (SRR) Scavenger Receptor Cysteine-rich; (GSL1) Gibberellin Stimulated-Like protein 1 (Callose synthase); (VLE1) Vegetative Lytic Enzyme 1; (GAS28) Gamete-specific 28 ***(D)*** Working model of CHT7-meidated cell wall regulation in Chlamydomonas.

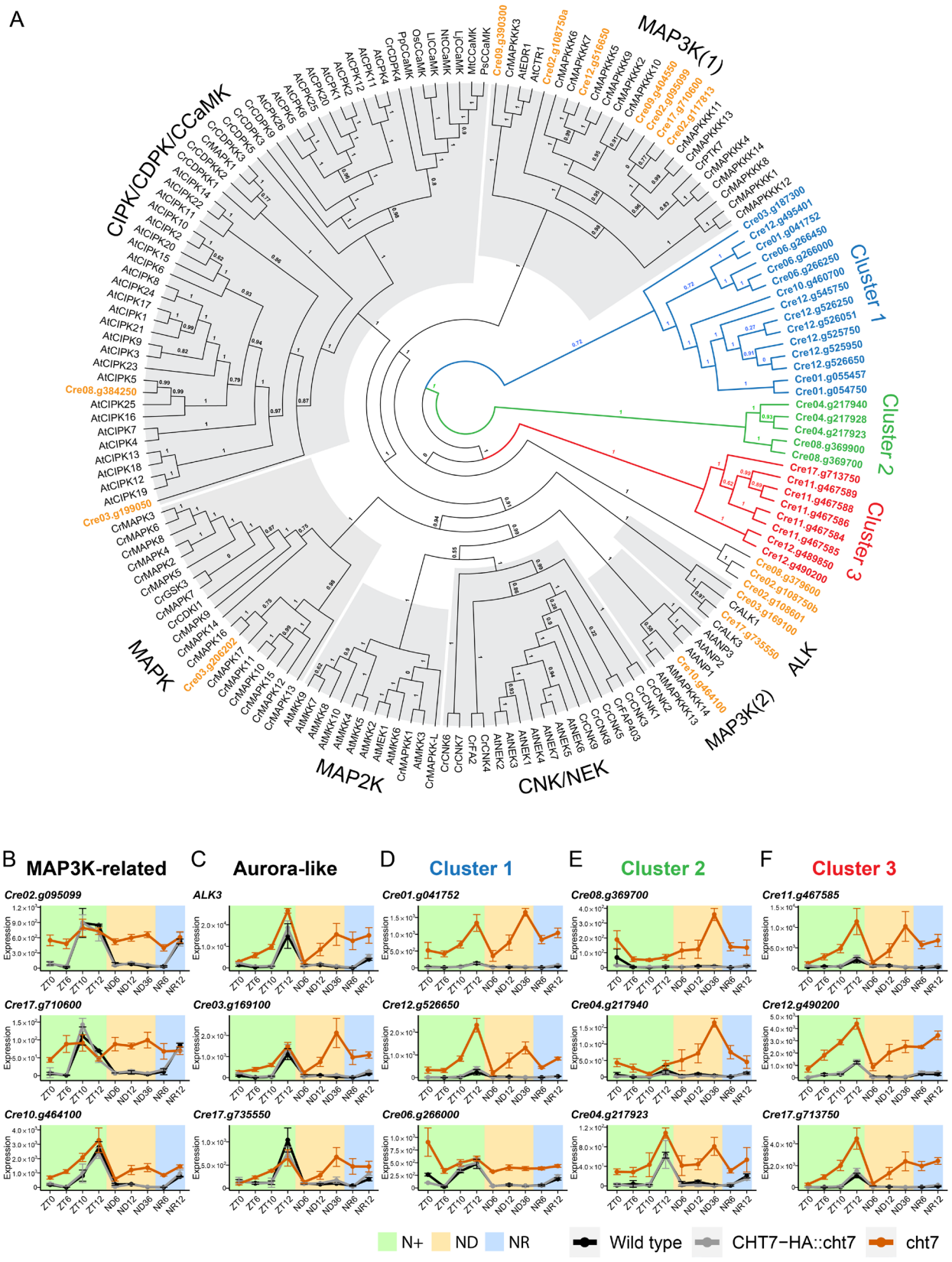


Figure 2.8. Analysis of protein kinase families misregulated in cht7.

Figure 2.8 (cont'd)

(A) Maximum likelihood phylogeny of the kinase domain of 43 protein kinases encoded by genes that are differentially expressed ($\log_2$ fold change > 1, *padj* < 0.05) in *cht7* at ≥ 5 timepoints along with kinase domains retrieved from 122 known protein kinases. Confidence of branches was tested by approximate Likelihood-Ratio (aLRT) branch support. Refer to Materials and Methods for details. (Cr) *Chlamydomonas reinhardtii*; (At) Arabidopsis; (Lt) *Lotus japonicus*; (Mt) *Medicago truncatula*; (Ps) *Pisum sativum*; (Nt) *Nicotiana tabacum*; (Ll) *Lilium longiflorum*; (Os) *Oryza sativa*; (Pp) *Physcomitrella patience*. (CCaMK) Ca2+/CaM-dependent Protein Kinase; (CIPK) CBL-interacting Protein Kinase; (CDK and CPK) Calcium-dependent Protein Kinases; (CNK) Chlamydomonas NIMA-related Kinase; (NEK) NIMA-related Kinase; (GSK3) Glycogen Synthase Kinase 3. ***(B-F)*** Expression of genes encoding MAPK3K-related kinases ***(B)***, Aurora-like kinases ***(C)***, and cluster 1-3 potential protein kinases ***(D-F)*** in *cht7* (orange), WT (black), and *CHT7-HA::cht7* (gray) across all timepoints. *Y*-axis: normalized read count.

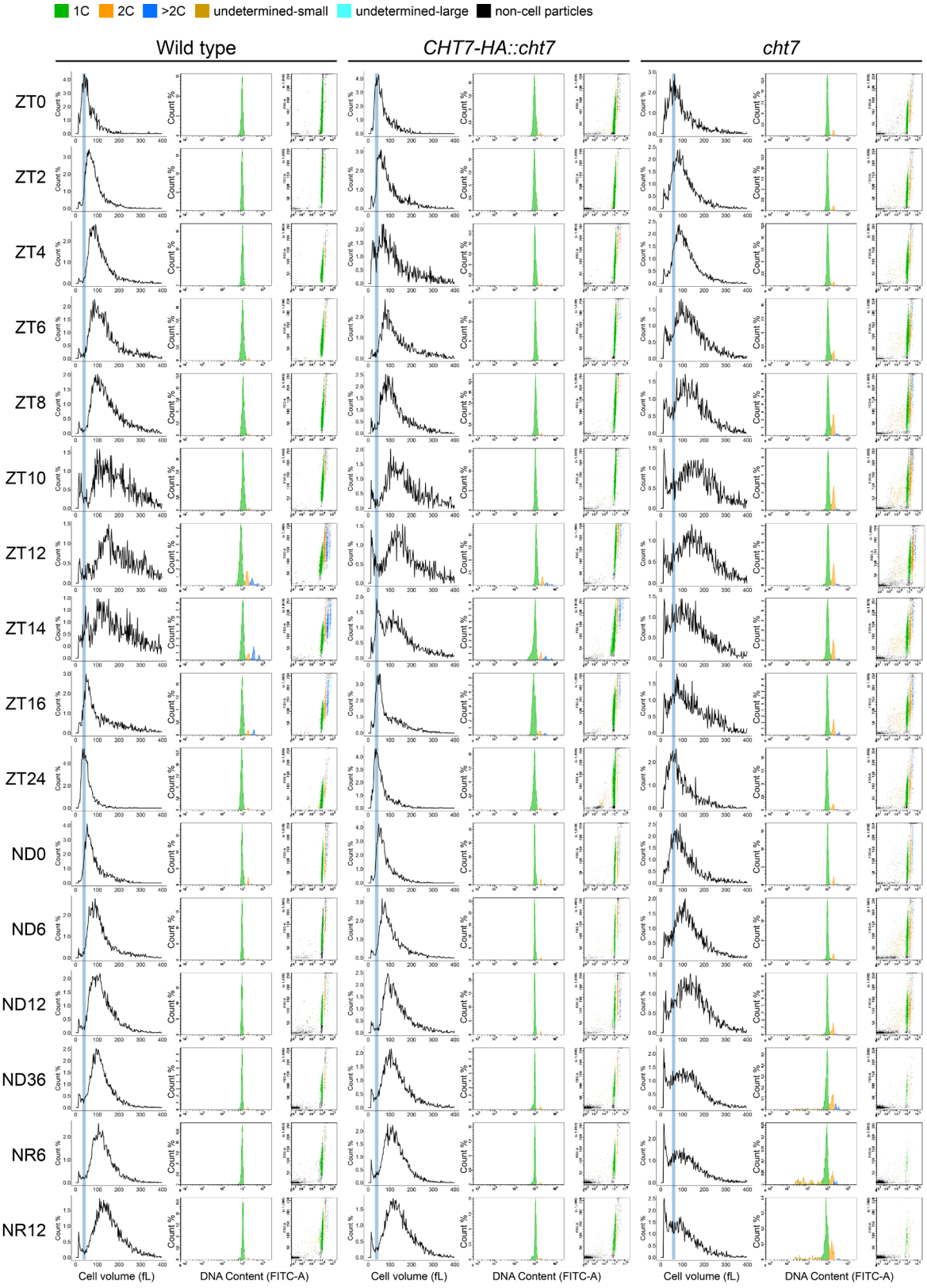


Figure S2.1. Cell cycle analysis of cell cultures grown under synchronized conditions.

Figure S2.1 (cont'd)

Cell size distribution (left column) and flow cytometry DNA analysis (middle and right columns) of synchronized cells of WT, *CHT7-HA::cht7*, and *cht7* grown under N-replete, N-deprived, and N-resupplied conditions. For cell size measurement (left column), cells sampled at different timepoints were fixed with 0.2% of Glutaraldehyde prior to analysis. *X*-axis: cell volume (fL). *Y*-axis: population (%) of cells. The most abundant population of the ZT0 sample is labeled in blue and serves as a reference for other samples. For DNA measurement (middle and right column), cells were fixed with ethanol: acetic acid fixation buffer and stained with DNA dye SYTOX Green, as described in Materials and Method. Middle column: histogram of cells stained with SYTOX Green. *X*-axis: DNA content determined by the intensity of FITC-A. *Y*-axis: population (%) of cells. Right column: scatter plot of cells stained with SYTOX Green. *X*-axis: DNA content determined by the intensity of FITC-A. *Y*-axis: Particle size determined by FSC-A (Forward scatter). Supplemental information to Figure 2.1.

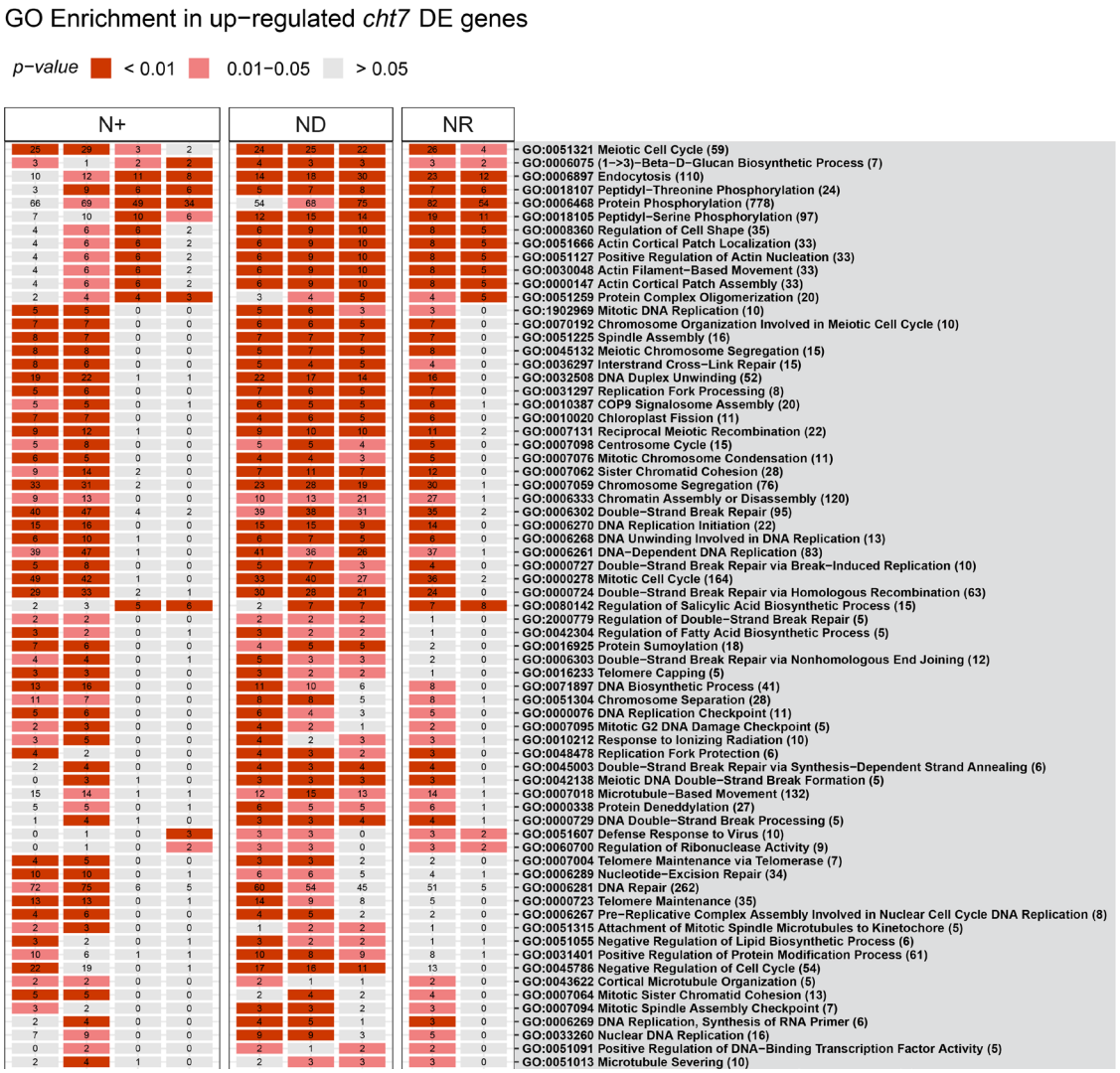
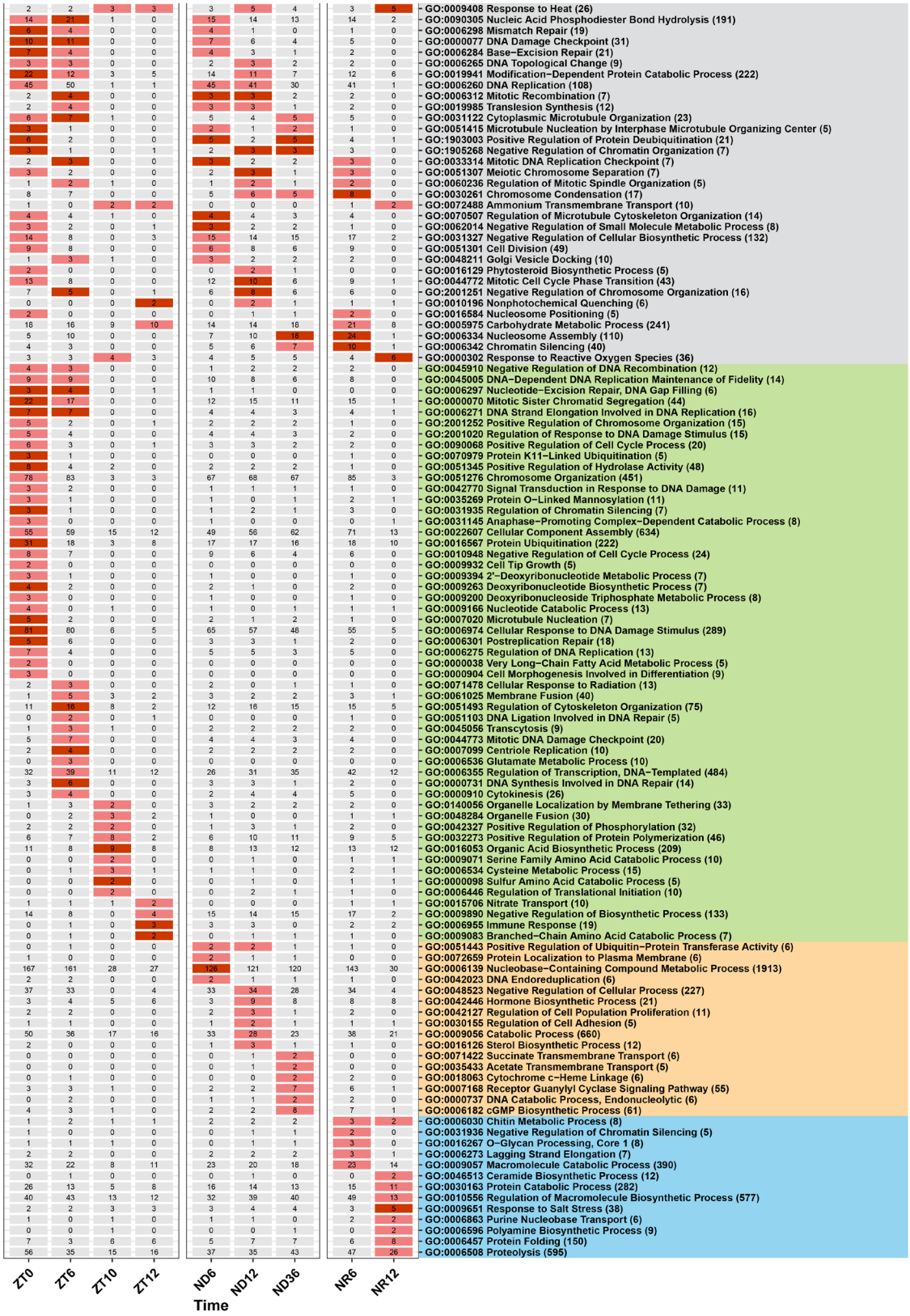


Figure S2.2. GO enrichment analysis of upregulated *cht7*-DEGs across all timepoints.

A total of 184 GO terms of biological processes were identified to be significantly ($P < 0.05$) enriched in the up-regulated *cht7*-DEGs at ≥ 1 timepoints. GO terms of biological processes that were explicitly enriched in N+ (light green), ND (light orange), NR (light blue) conditions are highlighted, and others (cross-conditions) are in grey. The number of up-regulated DEGs is labeled in the boxes. The number of total annotated genes is indicated after GO terms. Supplemental information to Figure 2.3.

Figure S2.2 (cont'd)



REFERENCES

REFERENCES

Alexa, A., and Rahnenfuhrer, J. (2020). topGO: Enrichment analysis for Gene Ontology. R package version 2.42.0.

Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22:1600-1607.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 55:539-552.

Atkins, K.C., and Cross, F.R. (2018). Interregulation of CDKA/CDK1 and the plant-specific cyclin-dependent kinase CDKB in control of the Chlamydomonas cell cycle. Plant Cell 30:429-446.

Beall, E.L., Manak, J.R., Zhou, S., Bell, M., Lipsick, J.S., and Botchan, M.R. (2002). Role for a Drosophila Myb-containing protein complex in site-specific DNA replication. Nature 420:833-837.

Bisova, K., Krylov, D.M., and Umen, J.G. (2005). Genome-wide annotation and expression profiling of cell cycle regulatory genes in *Chlamydomonas reinhardtii*. Plant Physiol 137:475-491.

Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics 31:3997-3999.

Bradley, B.A., and Quarmby, L.M. (2005). A NIMA-related kinase, Cnk2p, regulates both flagellar length and cell size in Chlamydomonas. J Cell Sci 118:3317-3326.

Bradley, B.A., Wagner, J.J., and Quarmby, L.M. (2004). Identification and sequence analysis of six new members of the NIMA-related kinase family in Chlamydomonas. J Eukaryot Microbiol 51:66-72.

Chan, K.L., Gligoris, T., Upcher, W., Kato, Y., Shirahige, K., Nasmyth, K., and Beckouet, F. (2013). Pds5 promotes and protects cohesin acetylation. Proc Natl Acad Sci U S A 110:13020-13025.

Cheeseman, I.M. (2014). The kinetochore. Cold Spring Harb Perspect Biol 6:a015826.

Chen, C., MacCready, J.S., Ducat, D.C., and Osteryoung, K.W. (2018). The molecular machinery of chloroplast division. Plant Physiol 176:138-151.

Chen, X.Y., and Kim, J.Y. (2009). Callose synthesis in higher plants. Plant Signal Behav 4:489-492.

Chen, X.Y., Liu, L., Lee, E., Han, X., Rim, Y., Chu, H., Kim, S.W., Sack, F., and Kim, J.Y. (2009). The Arabidopsis callose synthase gene GSL8 is required for cytokinesis and cell patterning. Plant Physiol 150:105-113.

Cheng, S.H., Willmann, M.R., Chen, H.C., and Sheen, J. (2002). Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. Plant Physiol 129:469-485.

Cronmiller, E., Toor, D., Shao, N.C., Kariyawasam, T., Wang, M.H., and Lee, J.H. (2019). Cell wall integrity signaling regulates cell wall-related gene expression in *Chlamydomonas reinhardtii*. Sci Rep 9:12204.

Cross, F.R. (2020). Regulation of multiple fission and cell-cycle-dependent gene expression by CDKA1 and the RB-E2F pathway in Chlamydomonas. Curr Biol 30:1855-1865 e1854.

Cross, F.R., and Umen, J.G. (2015). The Chlamydomonas cell cycle. Plant J 82:370-392.

De Benedetti, A. (2012). The Tousled-Like Kinases as Guardians of Genome Integrity. ISRN Mol Biol 2012:627596.

Doczi, R., Okresz, L., Romero, A.E., Paccanaro, A., and Bogre, L. (2012). Exploring the evolutionary path of plant MAPK networks. Trends Plant Sci 17:518-525.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 4:1184-1191.

Fabry, S., Muller, K., Lindauer, A., Park, P.B., Cornelius, T., and Schmitt, R. (1995). The organization structure and regulatory elements of Chlamydomonas histone genes reveal features linking plant and animal genes. Curr Genet 28:333-345.

Fang, S.C., de los Reyes, C., and Umen, J.G. (2006). Cell size checkpoint control by the retinoblastoma tumor suppressor pathway. PLoS Genet 2:e167.

Fei, X., Yu, J., Li, Y., and Deng, X. (2017). CrMAPK3 regulates the expression of iron-deficiency-responsive genes in *Chlamydomonas reinhardtii*. BMC Biochem 18:6.

Godin, S.K., Sullivan, M.R., and Bernstein, K.A. (2016). Novel insights into RAD51 activity and regulation during homologous recombination and DNA replication. Biochem Cell Biol 94:407-418.

Gray, J.V., Petsko, G.A., Johnston, G.C., Ringe, D., Singer, R.A., and Werner-Washburne, M. (2004). "Sleeping beauty": quiescence in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev 68:187-206.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307-321.

Hirano, T. (2012). Condensins: universal organizers of chromosomes with diverse functions. Genes Dev 26:1659-1678.

Hoffmann, X.K., and Beck, C.F. (2005). Mating-induced shedding of cell walls, removal of walls from vegetative cells, and osmotic stress induce presumed cell wall genes in Chlamydomonas. Plant Physiol 139:999-1014.

Horard, B., Sapey-Triomphe, L., Bonnefoy, E., and Loppin, B. (2018). ASF1 is required to load histones on the HIRA complex in preparation of paternal chromatin assembly at fertilization. Epigenetics Chromatin 11:19.

Joo, S., Nishimura, Y., Cronmiller, E., Hong, R.H., Kariyawasam, T., Wang, M.H., Shao, N.C., El Akkad, S.E., Suzuki, T., Higashiyama, T., et al. (2017). Gene regulatory networks for the haploid-to-diploid transition of *Chlamydomonas reinhardtii*. Plant Physiol 175:314-332.

Krysan, P.J., Jester, P.J., Gottwald, J.R., and Sussman, M.R. (2002). An Arabidopsis mitogen-activated protein kinase kinase kinase gene family encodes essential positive regulators of cytokinesis. Plant Cell 14:1109-1120.

Lang, W.C., and Chrispeels, M.J. (1976). Biosynthesis and release of cell wall-like glycoproteins during the vegetative cell cycle of *Chlamydomonas reinhardii*. Planta 129:183-189.

Li, Y., Liu, D., Lopez-Paz, C., Olson, B.J., and Umen, J.G. (2016). A new class of cyclin dependent kinase in Chlamydomonas is required for coupling cell size to cell division. Elife 5:e10767.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

Lucker, B.F., Hall, C.C., Zegarac, R., and Kramer, D.M. (2014). The environmental photobioreactor (ePBR): An algal culturing platform for simulating dynamic natural environments. Algal Res 6:242-249.

Luxmi, R., Blaby-Haas, C., Kumar, D., Rauniyar, N., King, S.M., Mains, R.E., and Eipper, B.A. (2018). Proteases shape the Chlamydomonas secretome: Comparison to classical neuropeptide processing machinery. Proteomes 6.

MacNeill, S.A. (2010). Structure and function of the GINS complex, a key component of the eukaryotic replisome. Biochem J 425:489-500.

Mahjoub, M.R., Qasim Rasi, M., and Quarmby, L.M. (2004). A NIMA-related kinase, Fa2p, localizes to a novel site in the proximal cilia of Chlamydomonas and mouse kidney cells. Mol Biol Cell 15:5172-5186.

Makharashvili, N., and Paull, T.T. (2015). CtIP: A DNA damage response protein at the intersection of DNA metabolism. DNA Repair (Amst) 32:75-81.

Marceau, A.H., Felthousen, J.G., Goetsch, P.D., Iness, A.N., Lee, H.W., Tripathi, S.M., Strome, S., Litovchick, L., and Rubin, S.M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun 7:12301.

Matsuo, T., Kuramoto, H., Kumazaki, T., Mitsui, Y., and Takahashi, T. (2012). LIN54 harboring a mutation in CHC domain is localized to the cytoplasm and inhibits cell cycle progression. Cell Cycle 11:3227-3236.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., et al. (2010). Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol 154:1737-1752.

Musacchio, A., and Salmon, E.D. (2007). The spindle-assembly checkpoint in space and time. Nat Rev Mol Cell Biol 8:379-393.

Nimonkar, A.V., Genschel, J., Kinoshita, E., Polaczek, P., Campbell, J.L., Wyman, C., Modrich, P., and Kowalczykowski, S.C. (2011). BLM-DNA2-RPA-MRN and EXO1-BLM-RPA-MRN constitute two DNA end resection machineries for human DNA break repair. Genes Dev 25:350-362.

Nisa, M.U., Huang, Y., Benhamed, M., and Raynaud, C. (2019). The plant DNA damage response: Signaling pathways leading to growth inhibition and putative role in response to stress conditions. Front Plant Sci 10:653.

Olson, B.J., Oberholzer, M., Li, Y., Zones, J.M., Kohli, H.S., Bisova, K., Fang, S.C., Meisenhelder, J., Hunter, T., and Umen, J.G. (2010). Regulation of the Chlamydomonas cell cycle by a stable, chromatin-associated retinoblastoma tumor suppressor complex. Plant Cell 22:3331-3347.

Pan, J., Wang, Q., and Snell, W.J. (2004). An aurora kinase is essential for flagellar disassembly in Chlamydomonas. Dev. Cell 6:445-451.

Park, J.J., Wang, H., Gargouri, M., Deshpande, R.R., Skepper, J.N., Holguin, F.O., Juergens, M.T., Shachar-Hill, Y., Hicks, L.M., and Gang, D.R. (2015). The response of *Chlamydomonas reinhardtii* to nitrogen deprivation: a systems biology analysis. Plant J 81:611-624.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 14:417-419.

Peters, J.M., Tedeschi, A., and Schmitz, J. (2008). The cohesin complex and its roles in chromosome biology. Genes Dev 22:3089-3114.

Riera, A., Barbon, M., Noguchi, Y., Reuter, L.M., Schneider, S., and Speck, C. (2017). From structure to mechanism-understanding initiation of DNA replication. Genes Dev 31:1073-1088.

Sadasivam, S., and DeCaprio, J.A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. Nat Rev Cancer 13:585-595.

Salanga, M., and Van Winkle-Swift, K. (2020). Structure and function of the primary zygote wall of *Chlamydomonas Monoica*. Microsc. Microanal. 7:50-51.

Salazar-Roa, M., and Malumbres, M. (2017). Fueling the cell division cycle. Trends in Cell Biology 27:69-81.

Schmit, F., Cremer, S., and Gaubatz, S. (2009). LIN54 is an essential core subunit of the DREAM/LINC complex that binds to the cdc2 promoter in a sequence-specific manner. FEBS J 276:5703-5716.

Schmollinger, S., Muhlhaus, T., Boyle, N.R., Blaby, I.K., Casero, D., Mettler, T., Moseley, J.L., Kropat, J., Sommer, F., Strenkert, D., et al. (2014). Nitrogen-sparing mechanisms in Clamydomonas affect the transcriptome, the proteome, and photosynthetic metabolism. Plant Cell 26:1410-1435.

Sijacic, P., Wang, W., and Liu, Z. (2011). Recessive antimorphic alleles overcome functionally redundant loci to reveal TSO1 function in Arabidopsis flowers and meristems. PLoS Genet 7:e1002352.

Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res 4:1521.

Sueoka, N. (1960). Mitotic replication of deoxyribonucleic acid in *Chlamydomonas reinhardi*. Proc Natl Acad Sci U S A 46:83-91.

Sumper, M., Berg, E., Wenzl, S., and Godl, K. (1993). How a sex pheromone might act at a concentration below 10(-16) M. EMBO J 12:831-836.

Sundin, L.J., Guimaraes, G.J., and Deluca, J.G. (2011). The NDC80 complex proteins Nuf2 and Hec1 make distinct contributions to kinetochore-microtubule attachment in mitosis. Mol Biol Cell 22:759-768.

Takeuchi, T., Lin, Y.T., Fekaris, N., Umen, J., Sears, B.B., and Benning, C. (2020a). Modulation of CHT7 complexes during light/dark- and nitrogen-mediated life cycle transitions of Chlamydomonas. Plant Physiol 184:1762-1774.

Takeuchi, T., Sears, B.B., Lindeboom, C., Lin, Y.T., Fekaris, N., Zienkiewicz, K., Zienkiewicz, A., Poliner, E., and Benning, C. (2020b). Chlamydomonas CHT7 is required for an effective quiescent state by regulating nutrient-responsive cell cycle gene expression. Plant Cell 32:1240-1269.

Tam, L.W., Wilson, N.F., and Lefebvre, P.A. (2007). A CDK-related kinase regulates the length and assembly of flagella in Chlamydomonas. J Cell Biol 176:819-829.

Thomas, J.H., Ceol, C.J., Schwartz, H.T., and Horvitz, H.R. (2003). New genes that interact with lin-35 Rb to negatively regulate the let-60 ras pathway in *Caenorhabditis elegans*. Genetics 164:135-151.

Tian, J., and Kong, Z. (2019). The role of the augmin complex in establishing microtubule arrays. J Exp Bot 70:3035-3041.

Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L.H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A.S., Kawaguchi, M., et al. (2006). Deregulation of a Ca2+/calmodulin-dependent kinase leads to spontaneous nodule development. Nature 441:1153-1156.

Tsai, C.H., Warakanont, J., Takeuchi, T., Sears, B.B., Moellering, E.R., and Benning, C. (2014). The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in Chlamydomonas. Proc Natl Acad Sci U S A 111:15833-15838.

Tulin, F., and Cross, F.R. (2014). A microbial avenue to cell cycle control in the plant superkingdom. Plant Cell 26:4019-4038.

Tulin, F., and Cross, F.R. (2015). Cyclin-dependent kinase regulation of diurnal transcription in Chlamydomonas. Plant Cell 27:2727-2742.

Vignard, J., Siwiec, T., Chelysheva, L., Vrielynck, N., Gonord, F., Armstrong, S.J., Schlogelhofer, P., and Mercier, R. (2007). The interplay of RecA-related proteins and the MND1-HOP2 complex during meiosis in *Arabidopsis thaliana*. PLoS Genet 3:1894-1906.

Vigneault, F., Lachance, D., Cloutier, M., Pelletier, G., Levasseur, C., and Seguin, A. (2007). Members of the plant NIMA-related kinases are involved in organ development and vascularization in poplar, Arabidopsis and rice. Plant J 51:575-588.

von der Heyde, B., and Hallmann, A. (2020). Targeted migration of pherophorin-S indicates extensive extracellular matrix dynamics in *Volvox carteri*. Plant J 103:2301-2317.

Wang, W., Sijacic, P., Xu, P., Lian, H., and Liu, Z. (2018). Arabidopsis TSO1 and MYB3R1 form a regulatory module to coordinate cell proliferation with differentiation in shoot and root. Proc Natl Acad Sci U S A 115:E3045-E3054.

Wheeler, G.L., Miranda-Saavedra, D., and Barton, G.J. (2008). Genome analysis of the unicellular green alga *Chlamydomonas reinhardtii* indicates an ancient evolutionary origin for key pattern recognition and cell-signaling protein families. Genetics 179:193-197.

Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699.

Wilson, N.F., and Lefebvre, P.A. (2004). Regulation of flagellar assembly by glycogen synthase kinase 3 in *Chlamydomonas reinhardtii*. Eukaryot Cell 3:1307-1319.

Yu, Y., Xia, X., Yin, W., and Zhang, H. (2007). Comparative genomic analysis of CIPK gene family in Arabidopsis and Populus. Plant Growth Regul. 52:101-110.

Zheng, S., Villa, R., Wang, J., Feng, Y., Wang, J., Becker, P.B., and Ye, K. (2014). Structural basis of X chromosome DNA recognition by the MSL2 CXC domain during Drosophila dosage compensation. Genes Dev 28:2652-2662.

Zones, J.M., Blaby, I.K., Merchant, S.S., and Umen, J.G. (2015). High-resolution profiling of a synchronized diurnal transcriptome from *Chlamydomonas reinhardtii* reveals continuous cell and metabolic differentiation. Plant Cell 27:2743-2769.

CHAPTER 3
Discovery of two potentially regulatory motifs enriched in misregulated genes of the *cht7* mutant and a potential binding protein

ABSTRACT

Transcriptome profile of the loss-of-function mutant of COMPROMISED HYDROLYSIS OF TRIACYLGLYCEROLS 7 (CHT7) in *Chlamydomonas reinhardtii* indicated that CHT7 affects transcriptional programs of several biological processes, including cell division, cell-wall remodeling, and new potential protein kinase cascades. To better understand the mechanistic details of CHT7-mediated gene regulation, I conducted a computational motif search in predicted promoter regions of genes highly misregulated in *cht7*. As a result, I found the enrichment of two DNA motifs, Motif 1 (5'-AGCTGTCR-3') and Motif1-associated Motif 2 (5'-GGATCCA-3'), in the misregulated genes of *cht7*. Motif 1 and 2 exhibit unique positioning with regard to the transcription start site (TSS), which is likely linked to the de-repression of genes in the *cht7* mutant. In addition, Motif 1 resembles cis-regulatory elements known to be recognized by several TALE-class homeoproteins in animals. A comparison of the homeodomain of five homeoproteins between Chlamydomonas and other species indicated that several key residues responsible for DNA binding specificity are conserved in Chlamydomonas homeoproteins HDG1 and GSP1. Together, these data provide further information on the regulatory mechanism mediated by a hypothesized CHT7 complex. Furthermore, a potential involvement of homeodomain proteins in the cell cycle regulation of Chlamydomonas is suggested.

INTRODUCTION

Based on our recently published work on the transcriptional analysis of the *cht7* mutant, the CHT7 protein appears to be a gene regulatory factor globally turning down gene expression (Lin et al., 2022). The misregulated genes in *cht7*, also known as *cht7*-differentially expressed genes (*cht7*-DEGs), generally can be categorized into three major groups according to their biological functions, including 1) genes for DNA replication and mitosis during cell division, 2) genes for cell-wall synthesis and remodeling, and 3) genes encoding three previously uncharacterized protein kinase families which may represent signaling cascades that need to be further explored. Each of the three *cht7*-DEG groups exhibits a unique expression pattern in *cht7* and two control lines (wild type and *CHT7HA::cht7*) throughout the cell cycle. For example, genes related to cell division are strongly de-repressed at non-cell dividing timepoints such as ZT0, ZT6, ND6-36, and NR6, while genes associated with cell-wall functions were abnormally induced at ND6-36 in *cht7* (ZT = Zeitgeber Time under N-replete (N+) condition; ND = N-deprived condition; NR = N-resupplied condition) (Lin et al., 2022). These findings indicate the complexity of the gene regulatory network mediated by CHT7 in response to different growth conditions. However, the molecular mechanism behind this regulatory network remains unclear.

As mentioned in Chapter 1, the animal homolog of CHT7, LIN54, is found in the MULTI-VULVAL CLASS B (MuvB) core, an essential regulatory complex mediating cell cycle progression. Studies showed that the MuvB core recruits transcription factors (TFs) or TF-associated proteins such as retinoblastoma protein P130, E2F, DP, and B-Myb to the promoter of target genes to positively or negatively affect their expressions (Marceau et al., 2016; Sadasivam and DeCaprio, 2013; Sadasivam et al., 2012). The DNA binding specificity of MuvB plays a crucial role in this process. The crystal structure of the CXC domain of LIN54 and its bound DNA

duplex revealed that LIN54 interacts with a specific DNA motif 5'-TTYYRAA-3' named the CHR site which has been widely found in the promoter region of late-cell cycle genes such as Budding Uninhibited By Benzimidazoles 1 (BUB1) and Polymerase Delta 1 (POLD1) in mammals (Marceau et al., 2016; Muller et al., 2014; Song et al., 2009). These studies suggest that as a member of the CXC protein family, Chlamydomonas CHT7 might also function in a DNA sequence-specific manner to negatively regulate the expression of its direct or indirect gene targets. Lending support to this theory is the presence of the CHR site within the TSS -2,500 bp region of some misregulated genes in *cht7*, such as CYCB1 and CDKB1 (Takeuchi et al., 2020b).

To investigate the potential interaction between CHT7 and DNA duplexes, a series of Chromatin immunoprecipitation (ChIP) experiments using ND cells of *CHT7HA::cht7* has been conducted previously (Takeuchi et al., 2020b). However, ChIP-qPCR probing of CHR sites did not yield any positive signal. In a preliminary experiment of ChIP-seq no evidence of CHT7-DNA interaction was observed, while the positive control using histone H3 successfully detected H3-bound DNA fragments (Takeuchi et al., 2020b). These results implied that the CHT7 protein might not bind to DNA directly. This hypothesis was supported by the finding that a predicted helix-coil-helix domain of the CHT7 protein, the P3 domain, was required for CHT7 to affect gene expression in genetic complementation experiments, while the proposed DNA-binding CXC domain seemed to be dispensable (Takeuchi et al., 2020b). The negative result of ChIP-qPCR also suggested that CHT7 or its complex might possibly interact with DNA motifs other than CHR sites. Lastly, the unsuccessful attempt to identify CHT7-bound DNA could also be explained by the possibility that the binding of CHT7 to DNA targets is a transient process and is thus difficult to detect. Transient binding between TFs and DNA is frequently observed in gene regulatory networks of plants (Alvarez et al., 2021; Alvarez et al., 2020). In either case, identifying any TFs associated with

CHT7 will provide valuable mechanistic details of the hypothesized CHT7-mediated aspects of cell cycle regulation in Chlamydomonas, and one possible approach is to find DNA motifs enriched in the promoter of *cht7*-DEGs and subsequently to use those motifs to identify the responsible TF.

The DNA-binding homeoproteins encoded by homeobox genes have been widely studied ever since they were discovered in *Drosophila* (Lewis, 1978). Here, I briefly review the current understanding of homeoproteins and explain how this information could potentially contribute to our understanding of CHT7 function.

The DNA sequence of Homeobox genes contains a conserved homeobox region which encodes a 60 amino acid long helix-loop-helix-turn-helix protein motif named homeodomain (HD) (Holland, 2013). HD proteins, also known as homeoproteins, are widely found in both unicellular and multicellular eukaryotes in the plant and animal kingdoms. Most homeoproteins are found to bind DNA through their HD in a sequence-specific manner (Burglin and Affolter, 2016). In general, HD proteins can be divided into the TALE (three amino acid loop extension) superclass and non-TALE classes (Burglin and Affolter, 2016). The HD of TALE homeoproteins has three extra amino acids in the first loop between helix 1 and 2 (Bertolino et al., 1995). In animals, there are five TALE-HD classes, namely MKX, IRO, TGIF, MEIS, and PBC (Burglin, 1997). The first TALE-HD protein that was identified is human TGIF1 which was named based on its binding specificity to the Thymine-Guanine (TG) DNA motif and its repression of gene expression (Bertolino et al., 1995). Later studies showed that the HD of both TGIF and MEIS-class proteins recognize a 5'-CTGTCA-3' (or 5'-TGACAG-3' in the reverse complement) motif, and such binding specificities are dependent on the residues at positions 47, 50, and 54 of the HD (Berger et al., 2008; Imoto et al., 2001; Jolma et al., 2013; Laughon, 1991). Notably, the TGIF and MEIS

HD proteins of humans were often linked with neuronal development and cell proliferation; mutation of these homeoproteins often is associated with cancer (Geerts et al., 2005; Melhuish et al., 2001; Wotton and Taniguchi, 2018). In flowering plants, there are 2 TALE-HD classes, BELL and KNOX, and 12 non-TALE HD classes. The Arabidopsis BELL1 is crucial to ovule development, while the KNOX HD proteins in Arabidopsis like Shoot Meristemless (STM), Knotted-Like From Arabidopsis Thaliana 1/2/6 (KNAT 1/2/6) are required for shoot apical meristem development and maintenance in Arabidopsis (Belles-Boix et al., 2006; Mukherjee et al., 2009; Reiser et al., 1995; Scofield et al., 2014). The DNA binding specificity of plant BELL and KNOX HD proteins have yet to be determined.

In comparison to the highly divergent homeoprotein family in multicellular organisms (255 in Human and 110 in Arabidopsis), the unicellular microorganism Chlamydomonas has only five homeoproteins, including the KNOX-class Gamete-Specific Minus Molecule 1 (GSM1), BELL-related Gamete-Specific Plus Molecule 1 (GSP1) and HDG1, and non-TALE HDZ1 and OCP3 (Holland, 2013; Mukherjee et al., 2009). GSP1 and GSM1 have been well-studied with regard to their collaborative functions in the sexual cycle of Chlamydomonas. The Minus Dominance (MID)-dependent expression of *GSP1* and *GSM1* is turned on during gametogenesis, resulting in the cytosolic accumulation of GSP1 and GSM1 proteins in mating-type plus and minus gametes, respectively. During gamete fusion, GSP1 and GSM1 form a heterodimer and translocate into the nucleus to initiate the transcriptional program for zygote formation (Kurvari et al., 1998; Lee et al., 2008; Lin and Goodenough, 2007). A transcriptome study of wild-type and *gsp1* zygotes revealed that the GSP1-GSM1 heterodimer is likely to target the TGAC motif to induce genes for zygotic cell wall formation and meiosis and to down-regulate genes for vegetative cell wall biosynthesis during the early-zygote stage (Joo et al., 2017). Together, these findings point out

several features of the Chlamydomonas homeoproteins, such as 1) the importance of dimerization on subcellular-localization and DNA binding activity and 2) a preserved DNA binding specificity within the TALE-class HD proteins. However, these findings also raise follow-up questions such as: Are homeoproteins in Chlamydomonas responsible for turning off cell cycle genes during zygote formation? Do any protein(s) assist the nuclear transport of the GSP1-GSM1 heterodimer? Also, what are the functions of other HD proteins like HDG1, HDZ1, and OCP3? Are they involved in regulating the asexual life cycle of vegetative cells? Lastly, does the GSM1 protein also form a heterodimer with the GSP1 paralog HDG1?

In this investigation, I performed a thorough computational analysis to find potential cis-regulatory elements (CREs) localized near the transcription start site (TSS) of *cht7*-DEGs. I found two DNA motifs, namely Motif 1 and Motif2, that were highly enriched in *cht7*-DEGs and exhibited a unique distributional pattern around the TSS. I also showed that the consensus sequence of Motif 1 was similar to CREs recognized by animal TGIF and MEIS-classes homeoproteins. The conserved nature of homeoproteins of various species suggests that Chlamydomonas HDG1 and GSP1 might have HDs similar to those of TGIF and MEIS homeoproteins. In summary, these results revealed the existence of CHT7-related DNA motifs, which led to hypothesize a previously uncharacterized TALE-HD mediated pathway in Chlamydomonas.

RESULTS

Discovery of DNA motifs: the approach

This study was aimed at identifying potential cis-regulatory elements (CREs) located in promoters of genes targeted by the CHT7 complex. I hypothesized that the regulatory DNA motifs targeted by the CHT7 complex are correlated with the transcriptional dynamics of the cell cycle. Accordingly, I performed a motif search using my recently published transcriptome dataset from synchronized cell cultures of *cht7*, wild type (WT), and the complementation line *CHT7HA::cht7* under autotrophic growth (Lin et al., 2022). The workflow of this motif analysis is illustrated in Fig. 3.1 and will be further described in later sections. As an overview, I collected a set of highly differentially expressed genes in *cht7* (*cht7*-DEGs) and grouped them into gene clusters exhibiting unique expression patterns across all timepoints. The pre-processing of DEGs with clustering is based on the hypothesis that specific expression patterns are dependent on unique CREs, so the motif search of individual gene clusters would generate more representative DNA motifs. The control group was composed of genes that were either not misregulated or were differentially expressed at ≤ 2 timepoints. Next, I located the transcription start site (TSS) of the target and control genes and extracted the TSS +/-250 bp sequences. The resulting 500 bp sequences of targets and controls were used as the input for finding motifs that were significantly enriched in target genes compared to the control group. Some motifs of interest were further analyzed for their positional distribution, co-occurrence with other motifs, and sequence similarities to known CREs. Lastly, I did a bottom-up, genome-wide search for genes containing the motif of interest and analyzed their annotated functions and expression patterns in *cht7*.

Preparation of the input DNA sequences for motif discovery

To search for potential CREs enriched in the misregulated genes in *cht7*, I targeted genes that were highly misregulated in *cht7* at ≥ 3 timepoints under normal and N-deprived/resupplied conditions during synchronous growth (*cht7*-DEG, n = 1,575) (Lin et al., 2022), and other genes were defined as the control gene set (n = 16,454). The selected *cht7*-DEGs were subdivided into several gene groups by hierarchical gene clustering. Figure 3.2A shows the hierarchical relationship between *cht7*-DEGs and the resulting branch (cluster) number when cutting at different heights. To balance the gene number and uniqueness of the expression pattern of each gene cluster, I cut the hierarchical tree at the height = 1.6, which resulted in six gene clusters (Cluster 1 to 6; abbreviated as C1 - C6). To test the unity of each gene cluster, I calculated the correlation (scored as Q-score; min -1, max 1) of gene components to the average centerline, and the result showed that five of the six gene clusters had an average Q-score > 0.65 (C1 to C5; Fig. 3.2B). This result suggested that gene clustering effectively categorized most genes based on their expression profiles. As shown in Fig. 3.2C, genes of C3 and C4 were strongly affected when CHT7 was absent. For instance, C4 genes in WT and *CHT7HA::cht7* were typically expressed at ZT10-12 and NR12, which were timepoints of cell division under synchronous growth conditions. GO enrichment analysis of individual clusters showed that biological processes related to DNA replication are highly enriched in genes of C4, which is consistent with our previous finding (Table 3.1) (Lin et al., 2022). For genes of C3, enriched GO terms include GO:0080142 Regulation of Salicylic Acid Biosynthetic Process and GO:0010387 COP9 Signalosome Assembly, which has also been described previously (Table 3.1) (Lin et al., 2022). Notably, the first one has been shown to consist of genes encoding flagellar-associated proteins but not Salicylic acid signaling in Chlamydomonas (Lin et al., 2022). Lastly, there are fewer genes of other clusters which did not show the enrichment

of specific biological pathways nor a significant mis-regulation in *cht7*. Thus, I concluded that among all DEGs in *cht7*, genes of C3 and C4 are the most affected by CHT7 as one would expect for genes directly targeted by the CHT7 complex, and their TSS +/- 250 bp were further analyzed for the potential presence of CHT7-associated regulatory motifs.

In-parallel motif discovery with STREME and HOMER

I performed motif analyses of *cht7*-DEGs with two commonly used software packages, STREME (Sensitive, Thorough, Rapid, Enriched Motif Elicitation) and HOMER (Hypergeometric Optimization of Motif EnRichment). The former belongs to the MEME suite (MEME: Multiple Em for Motif Elicitation) package (Bailey, 2021; Bailey et al., 2015; Heinz et al., 2010). While using different algorithms, both software packages are designed to discover novel motifs that are relatively enriched in target sequences compared to control sequences of the same length. As shown in Tables 3.1 and 3.2, both STREME and HOMER discovered several motifs that were significantly enriched (*p-value* < 0.05) in C3 and C4. The enrichment of STREME-detected motifs was further validated with SEA (Simple Enrichment Analysis) of the MEME-suite. For HOMER analysis, genes carrying the enriched motifs were identified with the FindMotif function of the same program. (Table 3.2 and 3.3). To narrow down candidate motifs associated with the CHT7 complex, I focused on motifs enriched in the C4 genes in this study, as they exhibited a significant de-repressed expression in the *cht7* mutant. To ensure the motifs have a specific, non-random sequence and were relatively unique in the target genes, I also focused on motifs with lengths ≥ 10 bp and that have a greater target-to-control percentage ratio. As a result, motifs STREME-1 and HOMER-2 of the C4 group were selected for further analysis.

Characterization of motifs STREME-1 and HOMER-2

I found that STREME-1 (CAGCTGTCRV) and HOMER-2 (CGACAGCTGS, reverse complement: SCAGCTGTCG) share a similar sequence. In addition, both motifs were found to be significantly enriched in C4 genes with similar gene numbers (STREME-1: *P*-value= 2.4E-05; HOMER-2: *P*-value= 1.0E-44) (Fig. 3.3A and Table 3.2 and 3.3). This result implied that both STREME and HOMER might detect the same motif with a consensus sequence of 5'-AGCTGTCR-3' (R= C or T). I then compared the gene composition of STREME-1 and HOMER-2 in the C4 and control gene groups, and found that genes having either HOMER-2 or STREME-1 near the TSS largely overlap (Fig. 3.3B). Figure 3.3C provides three examples of DNA sequences containing the motif STREME-1, HOMER-2, or both. In either case, both motifs had a consensus sequence of 5'-AGCTGTCR-3'. The variance between the two motifs might be due to specific differences in the algorithm of the two programs.

While STREME-1 and HOMER-2 were enriched in the TSS +/- 250 bp region of C4 genes, we noticed that quite a few genes of the control group also contained these motifs (Fig. 3.3A). Therefore, we inspected the positional distributions of STREME1 and HOMER-2 within the TSS +/- 250 bp regions and compared them between the two gene sets. As shown in Figure 3.3D, both motifs in the control group were evenly distributed from the -250 to 250 bp range. However, the same motifs in C4 genes exhibited a strong local preference between 0 to 100 bp downstream of the TSS. This finding suggested that STREME-1 and HOMER-2 motifs might have potentially unique functions for the C4 genes but were likely random occurrences in the control gene set. Together, I concluded that STREME-1 and HOMER-2 are generally the same motifs and are from here on, they are collectively designated as Motif 1 (5'-AGCTGTCR-3').

A second motif is enriched upstream of Motif 1 in *cht7*-DEGs

While Motif 1 (5'-AGCTGTCR-3') was significantly enriched in the TSS +/- 250 bp region of *cht7*-DEGs, the finding of Motif 1 in control genes implied that Motif 1 itself might not be sufficient for a given *cht7*-DEG to become a unique target of the CHT7 complex. In addition, previous studies have shown that CXC proteins interact with multiple TFs to regulate the expression of different genes (Fischer and DeCaprio, 2015; Sadasivam and DeCaprio, 2013). To test the possibility of the co-occurrence of Motif 1 and other regulatory motifs, I performed a second motif enrichment analysis in the TSS +/- 250 bp sequences of the Motif 1-containing genes in C4 *cht7*-DEGs (n= 325) and the control group (n= 2432). As shown in Figure 3.4A, both STREME and HOMER programs detected the enrichment of Motif 2 (consensus 5'-GGATCCA-3', M2-STREME and M2-HOMER) near the TSS of C4 genes compared to the control group. Interestingly, Motif 2 tended to be located at -50 to -70 bp upstream of the TSS for both C4 and control genes containing Motif 1, implying that the upstream position of Motif 2 to the TSS may not influence the expression of CHT7-regulated genes (Fig. 3.4B).

Notably, the presence of Motif 2 in promoter regions of C4 genes is not Motif 1-dependent. In fact, the original motif search has identified the enrichment of Motif 2-like motifs in promoter regions of C4 genes regardless of the existence of Motif 1 (light red, Table 3.2 and 3.3), which is further confirmed by motif enrichment analysis of Motif 2 by AME (Analysis of Motif Enrichment) of the MEME-suite package (Fig. 3.4C). To further test the independence of Motif 2 from Motif 1, I performed the positional distribution of Motif 2 in TSS +/- 250 bp regions of C4 and control genes without Motif 1. The result revealed that when Motif 1 is not present, Motif 2 is mainly located at two positions (-76.5 to -91 bp and +50 bp to TSS) in C4 genes, while Motif 2 in the control genes still remains to be located at -73 bp to the TSS (Fig. 3.4D). It is possible that the

additional occurrence of Motif 2 downstream of the TSS (+50 bp) is related to the gene regulation mediated by CHT7. Together, both Motif 1 and 2 might be recognized by the CHT7 complex, and the presence of both motifs enhances the control of the CHT7 complex on its target genes.

Next, I characterized the positional relationship between the two motifs in genes containing both motifs in the TSS +/- 250 bp region. The result showed that compared to the control genes, Motif 2 is located upstream of Motif 1 in most C4 genes (n= 130 upstream vs. n= 29 downstream, Fig. 3.4E). In addition, Motif 2 is generally closer to Motif 1 in the C4 genes than in the control set (Fig. 3.4E). Hypergeometric distribution was used to test the statistical significance of all possible Motif 2 - Motif 1 arrangements in the C4 and control gene sets, and obtained the results indicate a 100-150 bp gap between Motif 2 and Motif 1 with its significant enrichment in the C4 genes but not in the control (Fig. 3.4F). Together, these findings suggested that the combination of Motif 2 (5'-GGATCCA-3') and Motif 1 (5'-AGCTGTCR-3') within the TSS +/- 500 bp region is a unique feature in the de-repressed genes of *cht7*, and both motifs might serve as CREs recognized by the CHT7 complex.

Genome-wide characterization of genes containing Motif1 and 2

To validate the above-mentioned motif discovery results and investigate whether Motifs 1 and 2 play a role in CHT7-mediated gene regulation, I searched for genes carrying Motif 1, Motif 2, and both across the Chlamydomonas genome by FIMO (FIMO: Find Individual Motif Occurrences) of the MEME-suite. As a result, I found that 5,838 or 1,192 transcripts contained Motif 1 by using the qualifying threshold p-value of 1.0E-04 or 1.0E-05, respectively (lower *p*-value = closer to the consensus sequence of Motif 1). Similarly, there were 2,545 (*p*-value < 1.0E-04) and 436 (*p*-value

< 1.0E-05) genes containing Motif 2, and there were 981 genes containing both Motif 1 and 2 (*p*-value< 1.0E-4) near the TSS across the genome.

Next, I performed a series of hypergeometric tests to determine whether *cht7*-DEGs (misregulated at ≥ 1 timepoint) were enriched in each gene set when considering the positional distribution of both motifs. Firstly, I found that *cht7*-DEGs were highly enriched in genes with Motif 1 located downstream of the TSS and mildly enriched in genes having Motif 2 located both sides of the TSS (Fig. 3.5A). Second, the enrichment of *cht7*-DEGs in Motif 1 genes was more significant when Motif 1 was located at the position +1 to +100 bp of the TSS (Fig. 3.5B). Thirdly, when taking Motif 2 into consideration, I found that *cht7*-DEGs had the highest enrichment when Motif 2 was located 101 to 150 bp upstream of Motif 1, and the opposite order (Motif 1-Motif 2) was much less favorable in *cht7*-DEGs (Fig. 3.5C).

To test whether Motif 1 and 2 are potentially subject to CHT7-mediated gene regulation, I examined the transcription profile of the gene sets above in WT, *CHT7HA::cht7*, and *cht7* (Lin et al., 2022). I found that when using a lower threshold (*p*-value= 1.0E-4) to determine the presence of Motif 1, the detected genes (Motif 1 only) had no difference in gene expression between *cht7* and the two controls (Fig. 3.5D). However, when the threshold was raised to *p*-value= 1.0E-5, genes with Motif 1 located TSS +1 to +100 had a mildly de-repressed expression in *cht7* (Fig. 3.5E). The same observation was made for Motif 2-only genes (Fig. 3.5F-G). Notably, while Motif 2 is preferentially located upstream of the TSS, the presence of Motif 2 downstream of the TSS seemed to affect gene expression in the *cht7* mutant (Fig. 3.5G). Last but not least, when the two motifs co-occurred in the order Motif 2-Motif 1 across the TSS with a 101-150 bp gap in between the two motifs, genes carrying them were strongly de-repressed in *cht7* (Fig. 3.5H, right panel).

Finally, I inspected the annotated function of genes carrying Motif 1 alone (TSS +1 to +100 bp) and Motif 2-Motif 1 (Distance 101 to 150 bp). Most of these genes were involved in cell division (mostly chromosome segregation), protein processing (protein degradation and peptide modification), cell wall remodeling, and flagellar assembly (Fig. 3.6A and 3.6B). Representative examples of Motif 1 (5'-AGCTGTR-3') in genes of different biological processes are shown in Figure 3.6C.

Together, these findings suggested that Motif 1, Motif 2, and the co-occurrence of both motifs might represent a set of two CREs recognized by CHT7-associated TFs that repress the expression of the respective genes. I hypothesize that Motif 1 and Motif 2 are each sufficient to be recognized by CHT7-interacting TFs, which partially represses the transcription of genes containing it. When the two motifs are present in a certain order (Motif 2 – Motif 1), the binding of TFs to the two motifs inhibits the transcription further. The interaction between TFs and Motif 1 and 2 is dependent on CHT7. Therefore, in *cht7*, the repressive TFs can no longer bind to their target genes, and genes with the two motifs are de-repressed (Fig. 3.6D).

Motif 1 resembles CREs recognized by TALE homeoproteins

Results from the two independent, inverse motif analyses (gene-to-motif) and genome-wide gene search (motif-to-gene) suggested that Motif 1 (consensus 5'-AGCTGTR-3') is enriched in the TSS +1-100 bp region of misregulated genes in *cht7*. Here, I used Motif 1 as the query for searching established CREs and their corresponding TFs in the eukaryote TF database by the program Tomtom of the MEME-suite package. I found that Motif 1 is highly similar to DNA motifs recognized by some TALE-class homeoproteins. As shown in Fig. 3.7A and Table3.3, the conserved TGT of Motif 1 is also present in CREs specific to TALE homeoproteins such as MEIS2,

TGIF1/2, and PKNOX1 (Bertolino et al., 1995; Jolma et al., 2013; Yang et al., 2000). To test whether any of the 5 homeoproteins annotated in Chlamydomonas is related to the TALE homeoproteins shown in Figure 3.7A, I performed a phylogenetic analysis of HDs (Interpro ID: IPR001356) from species including *Homo sapiens* (Hs), *Drosophila melanogaster* (Dm), *Arabidopsis thaliana* (At), *Acetabularia acetabulum* (Aa), and *Chlamydomonas reinhardtii* (Cr). Overall, the result was in agreement with a previous study showing that CrGSM1 of Chlamydomonas is related to TALE-KNOX, while CrOCP3 and CrHDZ1 form distinct Non-TALE clades (Fig. 3.7B) (Lee et al., 2008). My data also showed that CrGSP1 and CrHDG1 are likely paralogs forming a unique clade, and that none of the annotated homeoproteins in Chlamydomonas are closely related to the animal TALE-TGIF and TALE-MEIS proteins (Fig. 3.7B).

In addition to the phylogenetic relationship, I also compared the HD peptide sequence of the homeoproteins above. As mentioned previously, residues at positions 47, 50, and 54 of α-helix 3 determine the DNA binding specificity. For example, a unique combination of N47-I50-R54 in some animal HDs was shown to interact with 5'-CTGTCA-3' (Fig. 3.7C) (Berger et al., 2008; Laughon, 1991). Both Chlamydomonas GSM1 and HDG1 contained N47 and I50 but had a different residue at position 54 (GSM1: K; HDG1: V; group I and III, Fig. 3.7C). Nevertheless, the identical residues at positions 47 and 50 support the hypothesis that GSM1 and HDG1 might recognize a DNA motif similar to TGIF and MEIS in animals. On the other hand, GSP1 had a residue combination of D47-T50-A54, implying that GSP1 targets DNA motifs other than HSG1 even though they are phylogenetically closely related (Fig. 3.7B and 3.7C).

Finally, I examined the expression of homeobox genes throughout the cell cycle of Chlamydomonas. The data showed that *GSM1* was expressed at ND36 in both WT, *CHT7HA::cht7*,

and *cht7*, which was consistent with its proposed protein function during gametogenesis (Lee et al., 2008) (Fig. 3.7D). Unexpectedly, gene expression of *GSP1* was highly induced in both *cht7* and *CHT7HA::cht7* compared to WT, suggesting that the MID-mediated repression of *GSP1* might be impaired by *cht7* background mutations (Lin and Goodenough, 2007; Tsai et al., 2014) (Fig. 3.7D). Lastly, the gene encoding HDG1 was highly expressed at ZT0 and slightly induced at ND36 in WT and *CHT7HA::cht7* but not in *cht7*, implying that HDG1 might be subject to the CHT7-mediated gene regulatory network and may have a different function than its HD paralogs GSP1 and GSM1 (Fig. 3.7D).

DISCUSSION

Previous work from our group involving extensive RNA-seq analysis of the *cht7* mutant and control strains has shown that CHT7 acts as a global regulator which represses a wide range of genes involved in cell cycle progression under normal, ND, and NR conditions (Chapter 2) (Lin et al., 2022; Takeuchi et al., 2020a; Takeuchi et al., 2020b; Tsai et al., 2018; Tsai et al., 2014). In this chapter, I explored potential TFs that may associate with CHT7 to mediate the repression of the respective genes. Preliminary experiments employing ChIP and Co-IP coupled with LC/MS failed to detect any TFs or TF-binding sites with which CHT7 directly interacts. Here, I took a different approach by computationally analyzing the DNA sequence within the TSS +/- 250 bp range of *cht7*-DEGs to search for potential regulatory elements recognized by the CHT7 complex and to identify TF candidates that might bind to these regulatory elements. I used two different tools, STREME and HOMER, to search for the existence of potentially conserved DNA motifs (Fig. 3.1) (Bailey, 2021; Bailey et al., 2015; Heinz et al., 2010). Both programs detected the enrichment of Motif 1, an 8-bp DNA sequence of 5'-AGCTGTCR-3', in C4 of *cht7*-DEGs which

is abnormally expressed in G0 and G1 stages of the cell cycle (Fig. 3.2 and 3.3). I noticed that not all C4 genes contained Motif 1 even though they were misregulated in *cht7*, which indicates some of the de-repressed genes in *cht7* may be indirect targets of the CHT7 complex, thus carrying a different CRE.

In addition to Motif 1, I found that a second motif, Motif 2 (5'-GGATCCA-3'), is enriched in both the Motif1-free and Motif 1-containing *cht7*-DEGs (Fig. 3.4A and C). The relationship between the two motifs and how they were positionally related to TSS was tested with hypergeometric distribution (Fig. 3.4E-F and 3.5A-C), a method that has been used for similar prior analyses (Sudarsanam et al., 2002; Zhang et al., 2012).

In comparison to Motif 1 which was enriched at the TSS +1 to +100 bp region of *cht7*-DEGs, Motif 2 seemed to mainly locate at TSS -50 to -100 bp in both *cht7*-DEGs and control genes (Fig. 3.3D and 3.4B). However, when Motif 1 is absent, Motif 2 also occurred at the TSS +50 bp in some *cht7*-DEGs, which was shown to be related to CHT7-mediated gene regulation (Fig. 3.4D and 3.5G). Furthermore, Motif 2 seems to have an additive effect on Motif 1 to derepress genes in the *cht7* mutant (Fig. 3.5H). Accordingly, I hypothesized that CHT7 is likely to interact with at least two TFs that individually bind to Motif 1 and 2 respectively. In my working model, there are three situations causing the inhibition of genes by CHT7: 1) the binding of the first TF on Motif 1 which is 1-100 bp downstream of the TSS, 2) the binding of the second TF to Motif 2 located downstream of the TSS when Motif 1 is not present, and 3) the binding of both TFs to Motif 1 (downstream of the TSS) and Motif 2 (upstream of the TSS) when the two motifs are 101-150 bp apart. The third situation may further suppress gene expression. Therefore, genes containing Motif 2 and Motif 1 are more de-repressed than genes having Motif 1 alone in the *cht7* mutant (Fig. 3.5H and 3.6D).

Lastly, I searched for CREs similar to Motif 1 in several TF databases. I found that the sequence of Motif 1 was closely related to several CREs recognized by human TALE HD proteins such as TGIF 1, TGIF2, MEIS2, and PKNOX1 (Fig. 3.7A and Table 3.4). It has been shown that animal homologs of TGIF1 and TGIF2 are transcriptional repressors. Besides, both TGIFs in humans were reported to interact with histone deacetylase to repress genes for early brain development (Melhuish et al., 2001; Wotton and Taniguchi, 2018).

In Chlamydomonas, there are 3 TALE HD proteins: GSM1, GSP1, and HDG1. Previous studies have shown that genes encoding GSP1 and GSM1 are specifically expressed in gametes of mating-type plus (mt+) and minus (mt-), respectively. Upon gamete fusion, cytosolic GSP1 and GSM1 form a heterodimer in the nucleus to trigger the expression of genes for zygotic cell wall formation and commitment of meiosis (Joo et al., 2017; Lee et al., 2008). Therefore, GSM1 and GSP1 are unlikely the general repressors of the vegetative cell cycle, leaving HDG1 as the possible candidate to be involved in this process.

The transcript of *HDG1* was detected in ND-induced gamete cells of both mating types (Lee et al., 2008). In our transcriptome data, the gene encoding HDG1 was highly expressed at ZT0 (the G0 phase) and mildly induced at ND36 (quiescence and probably gametic phase) in WT and *CHT7HA::cht7*. However, its expression was lower in *cht7* at the same timepoints compared to the controls (Fig. 3.7D). This result implies that HDG1 might be involved in regulating events at the G0 phase of the normal cell division cycle and during ND-induced quiescence. The lower expression of the *HDG1* gene in *cht7* might also contribute to the abnormal cell cycle gene expression in *cht7.*

Protein sequence analysis provided additional evidence supporting the idea that homeoproteins could bind to Motif 1 of *cht7*-DEGs in Chlamydomonas. As mentioned previously,

the DNA binding specificity of TALE-class homeoproteins is mainly dependent on three residues of the HD, P47, P50, and P54 (Berger et al., 2008; Laughon, 1991; Lee et al., 2008). The combination of N47-I50-R54 in animal TGIF and MEIS2 classes allows their HDs to recognize the 5'-TGTC-3' core motif (Fig. 3.7A, 3.7C, and Table 3.4)(Berger et al., 2008; Jolma et al., 2013). The Chlamydomonas HDG1 has N47-I50-V54 in the HD, which is distinct from other HD paralogs but identical to the Arabidopsis BELL class HD involved in shoot apical meristem development as well as ovule formation (Fig. 3.7C) (Kanrar et al., 2006; Reiser et al., 1995). The similarity of the three key residues between HDG1, TGIFs, and MEIS supports the hypothesis that HDG1 likely recognizes the TGT core of Motif 1 in Chlamydomonas. Lastly, the HDG1 protein is composed of 2,445 amino acids, as it is significantly longer than the other two TALE homeoproteins GSM1 (935 AA) and GSP1 (1,028 AA), inferring that HDG1 might have other domains (e.g., protein-protein interaction) which might be necessary for its function.

In summary, this study has provided further details of the molecular function of CHT7 in gene regulation in this study. The data led to hypothesize the presence of a homeoprotein-mediated pathway in the cell cycle regulation of Chlamydomonas. To test this hypothesis, one could introduce mutations to Motif 1 and 2 in several *cht7*-DEGs and determine if the two motifs are required for CHT7-mediated gene regulation. A yeast-two-hybrid assay or Co-IP assays could examine the interactions between CHT7 and HDG1. Assays of chromatin immunoprecipitation will verify the DNA binding activity of the HDG1 protein and the CHT7 complex. Finally, a loss-of-function mutant strain of HDG1 should be investigated, assuming it is viable, to determine whether HDG1 has overlapping functions with CHT7 as one would expect if they are active in the same complex and affect similar cellular processes.

MATERIALS AND METHODS

Clustering of *cht7*-DEGs

The total transcripts in *Chlamydomonas reinhardtii* (n= 19,526, *C.reinhardtii* v5.6; Phytozome genome ID: 281) were separated into two gene sets: 1) genes differentially expressed in *cht7* compared to both wild type and the complementation line *CHT7HA::cht7* at ≥ 3 timepoints of the total 9 timepoints under nitrogen (N)-replete (N+), N-deprived (ND), and N-resupplied (NR) conditions, as reported previously (Lin et al., 2022), and 2) genes that were only misregulated at ≤ 2 timepoints in *cht7* of the same transcriptome data set. The former was named cht7-DEGs (n= 1,586), and the latter was named the control group (n= 17,940). Both gene sets were further processed to remove transcripts with incomplete, duplicated, and overlapping flanking sequences on both sides of the TSS, resulting in smaller gene numbers for both categories (*cht7*-DEGs: n= 1,571; controls: n= 16,454). Prior to gene clustering, normalized read counts of *cht7*-DEGs were scaled to Z-scores by the equation $x-\bar{x}/\ \sigma$, where x is the read count of a timepoint and $\bar{x}$ is the mean of all timepoints, σ is the standard deviation of the mean. Hierarchical clustering of the gene matrix based on the Z-score to gene matrix was performed in R with the Pearson correlation option. Genes with a similar expression pattern across all nine timepoints were close to each other in the resulting hierarchical tree. The tree was then cut at height = 1.6, resulting in 6 branches (gene clusters). Each gene cluster's centerline (average Z-score) was calculated to represent the main expression pattern.

GO enrichment analysis

The method of GO enrichment analysis is described in Chapter 2. The input gene set for GO enrichment analysis was composed of the 6 gene clusters of a total of 1,157 genes misregulated in

cht7 at ≥ 3 timepoints (Fig 3.2). The analysis was performed using the R package topGO 2.40.0 (Alexa and Rahnenfuhrer, 2020; Alexa et al., 2006) with the *nodeSize = 5* option to construct the GO tree prior to analysis.

Motif discovery and characterization

Prior to motif analysis, the transcription start site (TSS) +/- 250 bp sequences of target genes were extracted with the R package biomaRt v2.50.3 (Durinck et al., 2009). These 500 bp-long sequences were "cleaned" first to remove duplicated sequences and overlapping regions between genes. The TSS +/- 250 bp sequences of individual clusters of the target (*cht7*-DEGs, C1 to C6, n= 1575) and control (non-*cht7* DEGs, n= 16,454) gene sets were analyzed together with the command-line version of programs STREME of the MEME-suite v5.4.1 (https://meme-suite.org/meme/) and HOMER v4.1 (http://homer.ucsd.edu/homer/) on the MSU-HPCC (MSU High-Performance Computing Cluster at ICER (https://icer.msu.edu/)). The control gene set was used to determine the unique enrichment of a motif in the target genes. The results of STREME were further analyzed with SEA (Simple Enrichment Analysis) and FIMO (Find Individual Motif Occurrences) of the MEME-suite for validating the motif enrichment and obtaining motif sites, respectively. The result from HOMER was also analyzed again with the HOMER findMotifs.pl -find <motif file> function to get motif sites. Motifs STREME-1 and HOMER-2 (collectively referred to as Motif 1) found in C4 genes were selected for further analysis. For the second motif analysis, the TSS +/- 250 bp sequences of the target (Motif 1-containing Cluster 4 of *cht7*-DEGs, n= 369) and control (Motif 1- containing non-*cht7* DEGs, n= 2,917) gene sets were processed using the same procedure as above. The processing of the resulting data, including gene analysis, motif patterns, and statistical tests, were performed with R v4.0 on RStudio (v1.3.959). AME (Analysis of Motif Enrichment)

of the MEME-suite was used to analyze the enrichment of Motif 2 in TSS +/- 250 bp regions of genes that do not contain Motif 1 in the same region. The positional distribution of Motif 1 and 2 was performed on CenriMo of the MEME-suite using the position probability matrix of each motif and the corresponding TSS +/- 250 bp sequences as the input. Lastly, a sequence comparison of Motif 1 with known motifs was performed using Tomtom of the MEME-suite against databases including JASPAR 2020 (https://jaspar.genereg.net/) and Arabidopsis DAP (http://neomorph.salk.edu/dap_web/pages/index.php).

Statistical analysis of motif co-occurrence

A *cht7*-DEG was defined as any gene being misregulated in *cht7* compared to the control at ≥1 timepoint during the synchronous growth, as described in Lin et al., 2022. The use of cumulative hypergeometric distribution to assess the significance of motif co-occurrence in a given gene set was reported previously (Sudarsanam et al., 2002; Zhang et al., 2012). For testing the enrichment of co-occurrence of M1 and M2 in *cht7*-DEGs and control genes, I used the following equation:

$$P(q) = \sum_{q}^{m} \frac{\frac{m}{q}\frac{n}{k-q}}{\frac{m+n}{k}}$$

Where *P* is the probability of getting *q* number of *cht7*-DEGs or control genes with co-occurrence of M1 and M2. *m* is the total number of genes with a co-occurrence of M1 and M2. *n* is the number of genes without co-occurrence of M1 and M2. *k* is the gene number of *cht7*-DEG or the control group.

To test the enrichment of *cht7*-DEGs or control genes in genes containing various motif combinations, I used the same equation above, where *q* is the number of *cht7*-DEGs or control genes with a given motif combination. *m* is the number of *cht7*-DEGs or control genes. *n* is the

number of non-*cht7* DEGs or non-control genes. *k* is the number of genes with a given motif combination. The *p*-value is computed using the *phyper* function in R with the *lower.tail = TRUE* and *log.p = TRUE* functions.

Multiple sequence alignment and phylogenetic analysis of homeoproteins

The full-length peptide sequences of the five Chlamydomonas homeoproteins were obtained from Phytozome v12.1 (dataset: Creinhardtii_281) using R package biomaRt 2.44.0 (Durinck et al., 2009). Sequences of the 20 selected homeoproteins from *Arabidopsis thaliana*, *Homo sapiens*, *Drosophila melanogaster*, and *Acetabularia acetabulum* were obtained from the Uniprot database (https://www.uniprot.org/). Sequences of the HD of all homeoproteins above were retrieved by searching for IPR001356 Homeobox domain in InterPro (https://www.ebi.ac.uk/interpro/). Multiple sequence alignment of the 25 HDs was performed using ClustalOmega in the R package msa 1.19.0 (Bodenhofer et al., 2015). Maximum likelihood (ML) phylogeny was estimated using the Whelan and Goldman (WAG) substitution model implemented in the PhyML 3.3 software (Guindon et al., 2010; Whelan and Goldman, 2001), and branch support was measured by a parametric, Chi2-based approximate Likelihood-Ratio (aLRT) test (Anisimova and Gascuel, 2006). The resulting unrooted ML tree was further analyzed with FigTree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). Lastly, the result of multiple sequence alignment was displayed using the python package pyBoxshade (https://github.com/mdbaron42/pyBoxshade)

APPENDIX

Cluster	GO term (Biological Process)	Hit	*P*-value
1	GO:0006145-Purine Nucleobase Catabolic Process	2/6	4.00E-04
	GO:0015707-Nitrite Transport	2/9	9.60E-04
	GO:0015706-Nitrate Transport	2/10	1.19E-03
	GO:0046083-Adenine Metabolic Process	1/5	2.62E-02
	GO:0000098-Sulfur Amino Acid Catabolic Process	1/5	2.62E-02
2	GO:0009768-Photosynthesis, Light Harvesting in Photosystem I	2/25	7.60E-04
	GO:0018298-Protein-Chromophore Linkage	2/26	8.20E-04
	GO:0009416-Response to Light Stimulus	2/85	8.56E-03
	GO:0009394-2'-Deoxyribonucleotide Metabolic Process	1/7	1.17E-02
	GO:0009263-Deoxyribonucleotide Biosynthetic Process	1/7	1.17E-02
3	GO:0080142-Regulation of Salicylic Acid Biosynthetic Process	7/15	2.10E-09
	GO:0007131-Reciprocal Meiotic Recombination	5/22	2.90E-05
	GO:0010387-COP9 Signalosome Assembly	4/20	3.20E-04
	GO:0060700-Regulation of Ribonuclease Activity	3/9	3.80E-04
	GO:0051607-Defense Response to Virus	3/10	5.40E-04
4	GO:0006270-DNA Replication Initiation	16/22	8.90E-18
	GO:0000724-Double-Strand Break Repair via Homologous Recombination	26/63	4.30E-11
	GO:0031297-Replication Fork Processing	7/8	2.40E-09
	GO:0070192-Chromosome Organization Involved in Meiotic Cell Cycle	7/10	3.30E-08
	GO:0032508-DNA Duplex Unwinding	19/52	5.00E-08
5	GO:0010196-Nonphotochemical Quenching	2/6	3.10E-04
	GO:0006897-Endocytosis	3/110	1.41E-02
	GO:0009179-Purine Ribonucleoside Diphosphate Metabolic Process	2/27	1.80E-02
	GO:0010506-Regulation of Autophagy	1/5	2.30E-02
	GO:0010016-Shoot System Morphogenesis	1/5	2.30E-02
6	GO:0034308-Primary Alcohol Metabolic Process	1/5	6.40E-03
	GO:0006897-Endocytosis	2/110	8.30E-03
	GO:0046185-Aldehyde Catabolic Process	1/7	9.00E-03
	GO:0007004-Telomere Maintenance via Telomerase	1/7	9.00E-03
	GO:0006740-NADPH Regeneration	1/15	1.92E-02

Table 3.1. GO Enrichment analysis of Cluster 1 to 6 of *cht7*-DEGs

A total of 1,571 transcripts differentially expressed in *cht7* compared to wild type and *CHT7-HA::cht7* at ≥ 3 timepointa was grouped into 6 gene clusters by hierarchical clustering based on their Z-scores, which were further analyzed for enriched GO terms. For each gene cluster, the top five enriched GO terms of Biological Process are listed. Hits (x/y): number of the input genes (x) over total genes (y) within a GO term.

					STREME					SEA					
Cluster	**Input/Ctrl**	**Motif_ID**	**Consensus**	**Width (bp)**	**TP**	**TP (%)**	**FP**	**FP (%)**	***P*-value**	**TP**	**TP (%)**	**FP**	**FP (%)**	***P*-value**	***E*-value**
3	403/16454	STREME-1	ACTATCCAGCAGCACCTT	18	9	2.2%	3	0.0%	5.5E-04	7	1.9%	3	0.0%	4.8E-10	8.6E-09
	403/16454	STREME-2	GCCGTAC	7	81	20.1%	1687	10.3%	2.3E-03	50	13.8%	1027	6.9%	4.6E-06	8.3E-05
	403/16454	STREME-3	CAACTGCC	8	215	53.3%	4840	29.4%	6.9E-03	126	34.7%	2404	16.2%	2.4E-17	4.3E-16
4	912/16454	STREME-1	CAGCTGTCRV	10	334	36.6%	2639	16.0%	2.4E-05	225	27.4%	1580	10.7%	7.7E-38	5.4E-37
	912/16454	STREME-2	GCTGGATCCADDCY	14	215	23.6%	1109	6.7%	9.4E-04	297	36.2%	2571	17.4%	1.1E-35	8.0E-35
	912/16454	STREME-3	GTAAAA	6	648	71.1%	10024	60.9%	4.6E-03	658	80.2%	10841	73.2%	4.0E-06	2.8E-05

Table 3.2. Enriched DNA motifs in Cluster 3 and 4 of *cht7*-DEGs discovered by STREME.

The enrichment of individual DNA motifs in *cht7*-DEGs was detected by the program STREME and validated by the software SEA of the MEME-suite package. TP (True positive): Number of motif-carrying *cht7*-DEGs (Input). FP (False positive): Number of motif-carrying control genes (Ctrl). *E*-value: adjusted p-value. Light green: Motif 1. Light red: Motif 2-like.

					HOMER					HOMER-FindMotif					
Cluster	Input/Ctrl	Motif ID	Consensus	Width (bp)	TP	TP (%)	FP	FP (%)	*P*-value	TP (sites)	TP (genes)	TP (%)	FP (sites)	FP (genes)	FP (%)
3	403/16454	HOMER-1	CAACTGCTGC	10	55	13.7%	468.7	2.9%	1.0E-19	56	55	13.6%	466	461	2.80%
	403/16454	HOMER-2	CCAACCGYW	9	97	24.1%	1462.1	8.9%	1.0E-18	110	97	24.1%	1524	1434	8.70%
	403/16454	HOMER-3	WGAGCCCGTTGC	12	17	4.2%	26.8	0.2%	1.0E-16	17	17	4.2%	27	27	0.20%
4	912/16454	HOMER-1	CGACAGCTG	9	288	31.7%	2119.4	12.9%	1.0E-45	337	289	31.7%	2316	2038	12.40%
	912/16454	HOMER-2	CGACAGCTGS	10	234	25.7%	1495.9	9.1%	1.0E-44	263	234	25.7%	1517	1410	8.60%
	912/16454	HOMER-3	TGGATCCAGN	10	224	24.6%	1414.8	8.6%	1.0E-43	386	224	24.6%	2014	1382	8.40%

Table 3.3. Enriched DNA motifs in Cluster 3 and 4 of *cht7*-DEGs discovered by HOMER.

The enrichment of individual DNA motifs in *cht7*-DEGs was detected by the program HOMER and validated by the FindMotif function of the same program. TP (True positive): Number of motif-carrying *cht7*-DEGs (Input). FP (False positive): Number of motif-carrying control genes (Ctrl). Light green: Motif 1. Light red: Motif 2-like.

TFs	*E* -value	Organisms	Database	Accessions	Concensus	Reference
TGIF1	2.10E-02	*Homo sapiens*	JASPAR 2020	MA0796.1	TGACAGCTGTCA	Jolma et al., 2013
TGIF2	2.93E-02	*Homo sapiens*	JASPAR 2020	MA0797.1	TGACAGsTGTCA	
PKNOX	8.67E-02	*Homo sapiens*	JASPAR 2020	MA0782.1	TGACAGsTGTCA	
MEIS2	1.80E-02	*Homo sapiens*	HumanTF 1.0	MEIS2_DBD_1 (HumanTF 1.0)	wTGACAGsTGTCAm	

Table 3.4. Established homeoprotein TFs and corresponding CREs.

Information of TFs which recognize CREs similar to Motif 1. Data was retrieved from JASPAR and HumanTF. The Tomtom-computed *E*-value (adjusted *P*-value) represents the probability of false-positive events which are matches between random motifs and target motifs. Lower *E*-value= Higher confidence of the match.

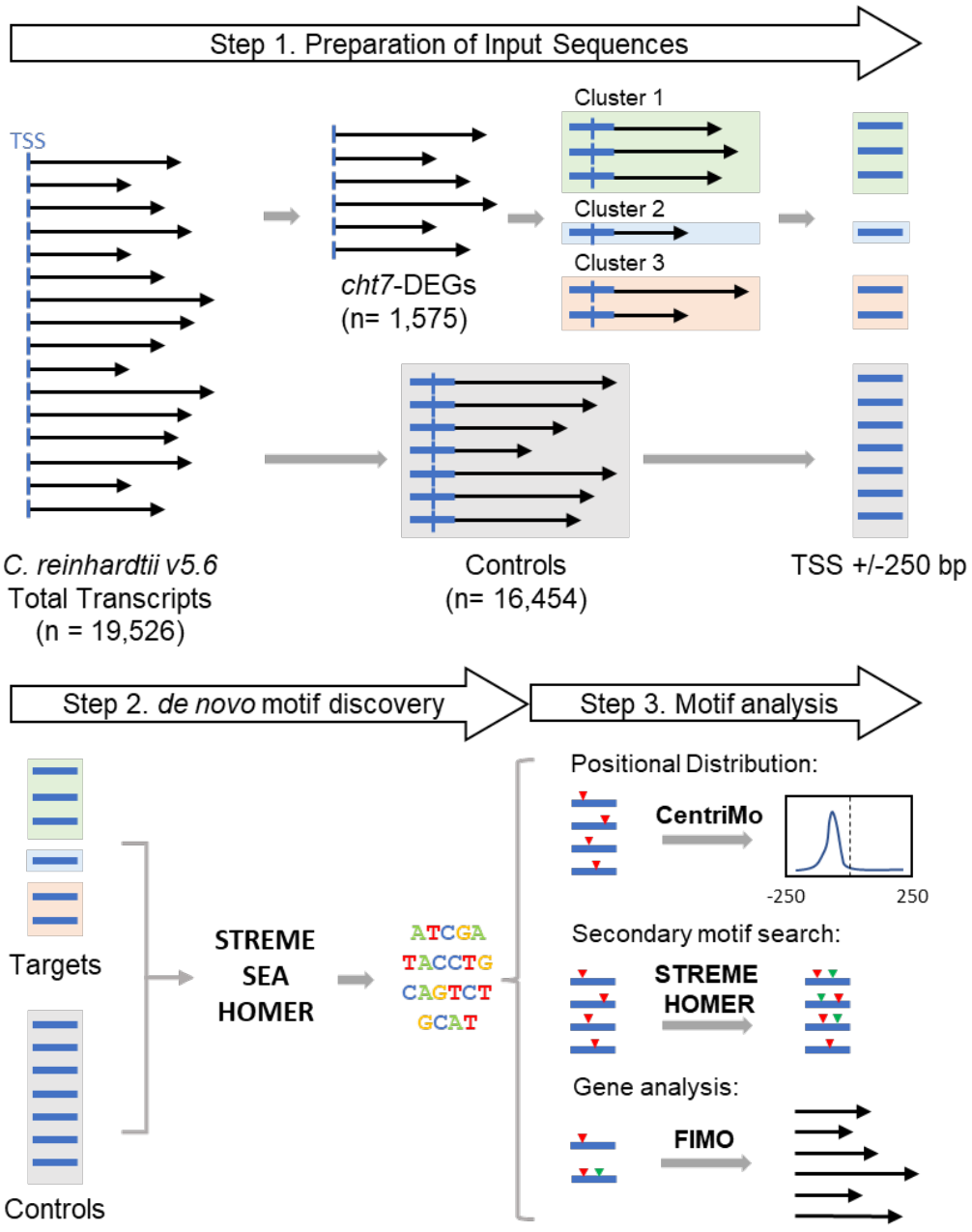


Figure 3.1. Discovery of regulatory elements in *cht7*-DEGs.

A workflow for motif finding in promoters of *cht7*-DEGs (DEG: differentially expressed genes). Step 1: Total transcripts (n= 19,526) of *Chlamydomonas reinhardtii* v5.6 were processed and separated into *cht7*-DEG (targets, n= 1,575) and non-*cht7* DEG (controls, n= 16,454) groups (see Materials and Methods for details of data processing). The former were further subdivided into sub-gene clusters based on the gene expression patterns at different timepoints under N+ (nitrogen-

Figure 3.1 (cont'd)

replete), ND (nitrogen-deprived), and NR (nitrogen-resupplied) conditions as described in Chapter 2 and Lin et al., 2022. The TSS +/- 250 bp sequences of individual transcripts in both groups were extracted as the input sequence for Step 2 (TSS: Transcription start site). Step 2: The target and control TSS +/- 250 bp sequences were analyzed for motif discovery with software from the MEME-suite (STREME and SEA) and HOMER to find DNA motifs relatively enriched in the target set comparted to the control set. Step 3: The discovered motifs were further analyzed for their positional preference within the TSS +/- 250 bp region by the CentriMo program of the MEME-suite. Some motifs of interest were selected for a second motif enrichment test on genes carrying these primary motifs. Motifs of interest were also used for a reverse search against the genome to find genes containing these motifs in the TSS +/- 250 bp region for further analyses.

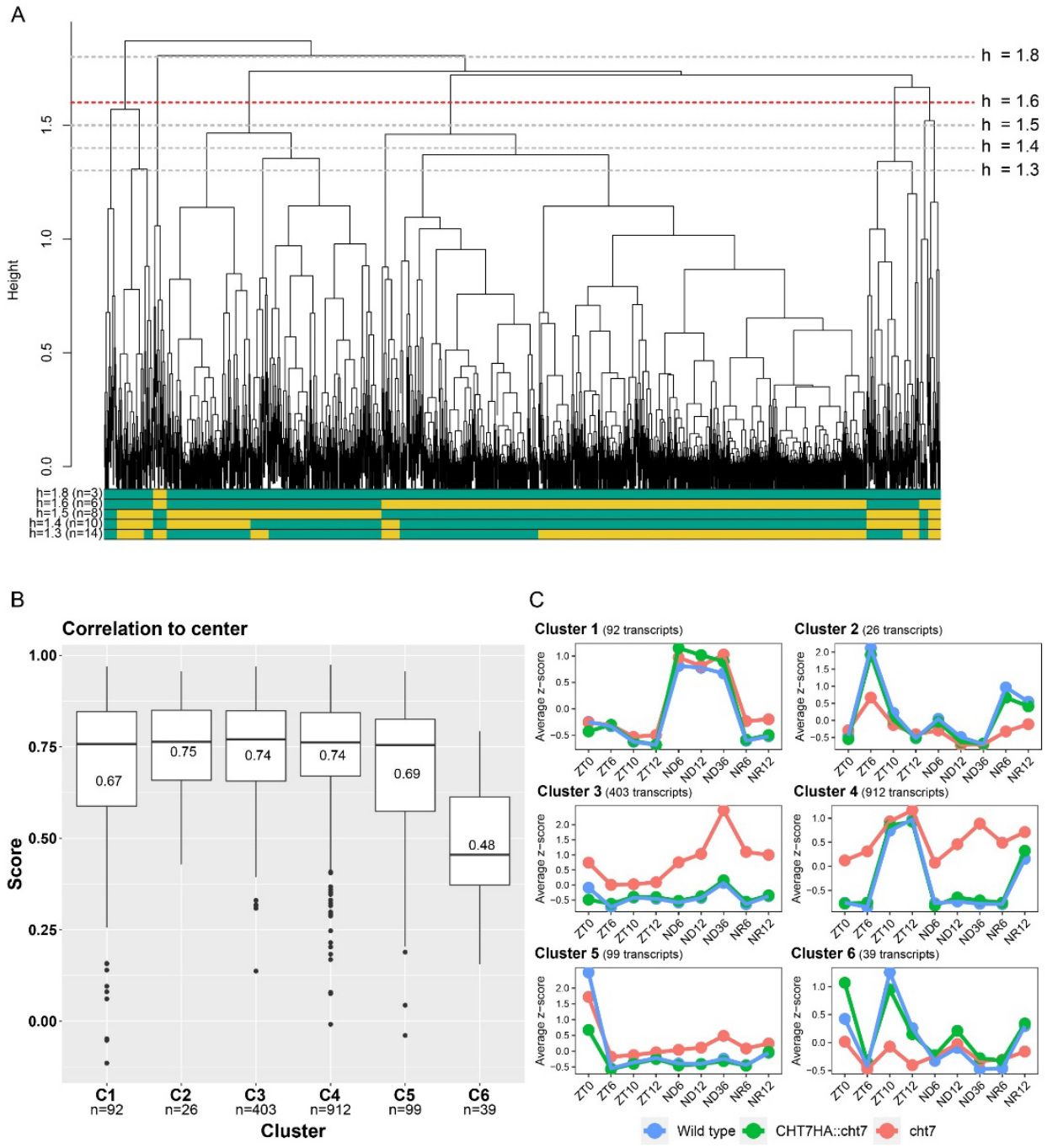


Figure 3.2. Hierarchical gene clustering of *cht7*-DEGs.

(A) Hierarchical tree illustrating the relationship between individual *cht7*-DEGs (n= 1,575) constructed based on their expression profiles under N+, ND, and NR growths. Dashed lines indicate the height of the tree, and the red line is the selected height for cutting the tree to generate

Figure 3.2 (cont'd)

gene clusters in this study. Color bars (green and yellow) represent gene clusters. ***(B)*** Boxplots of the gene-to-center correlation score of Cluster 1 to 6 after cutting the tree at height= 1.6. ***(C)*** Average gene expression (Z-score) of total genes in wild type (WT, blue), CHT7HA::cht7 (green), and cht7 (pink) in Cluster 1 to 6.

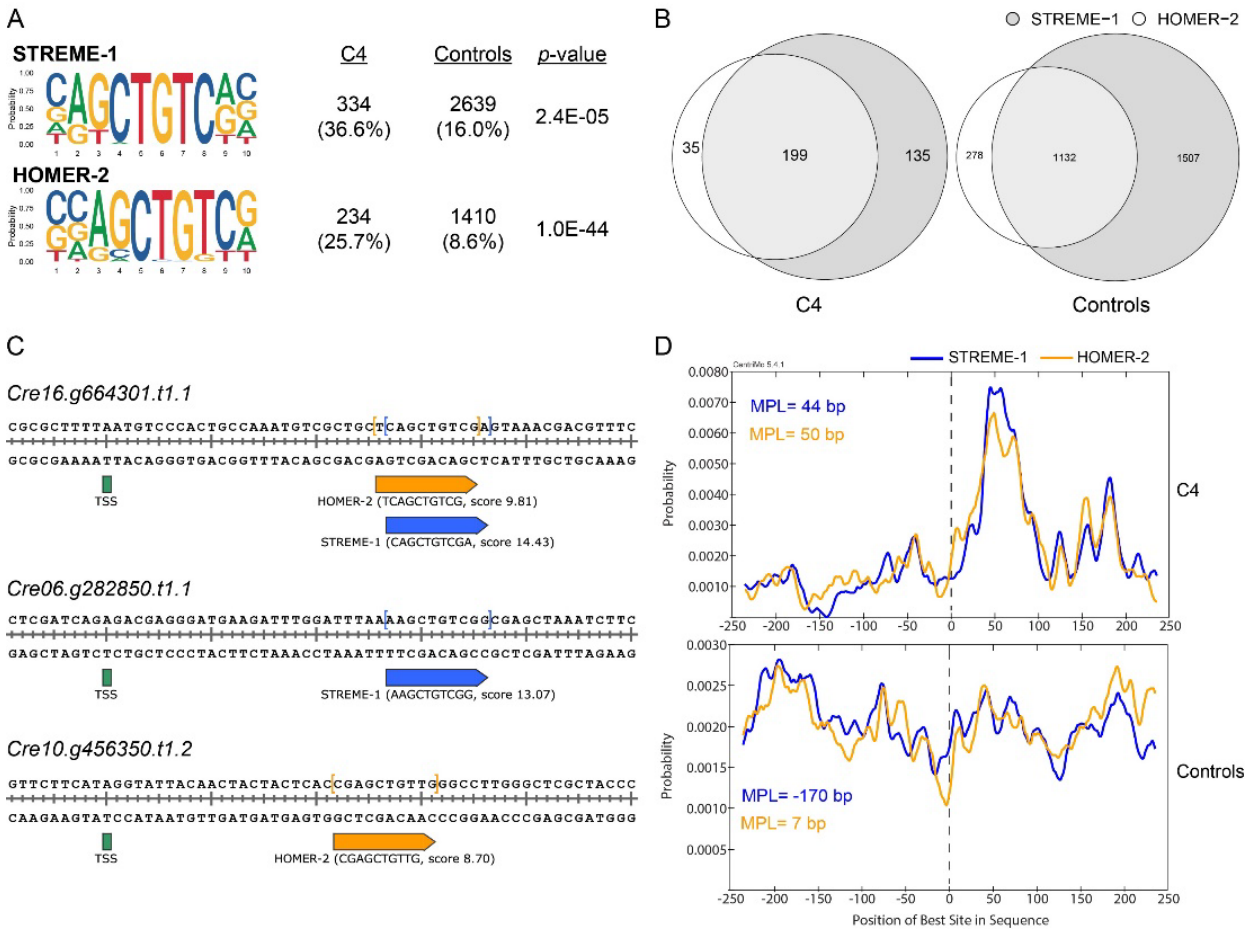
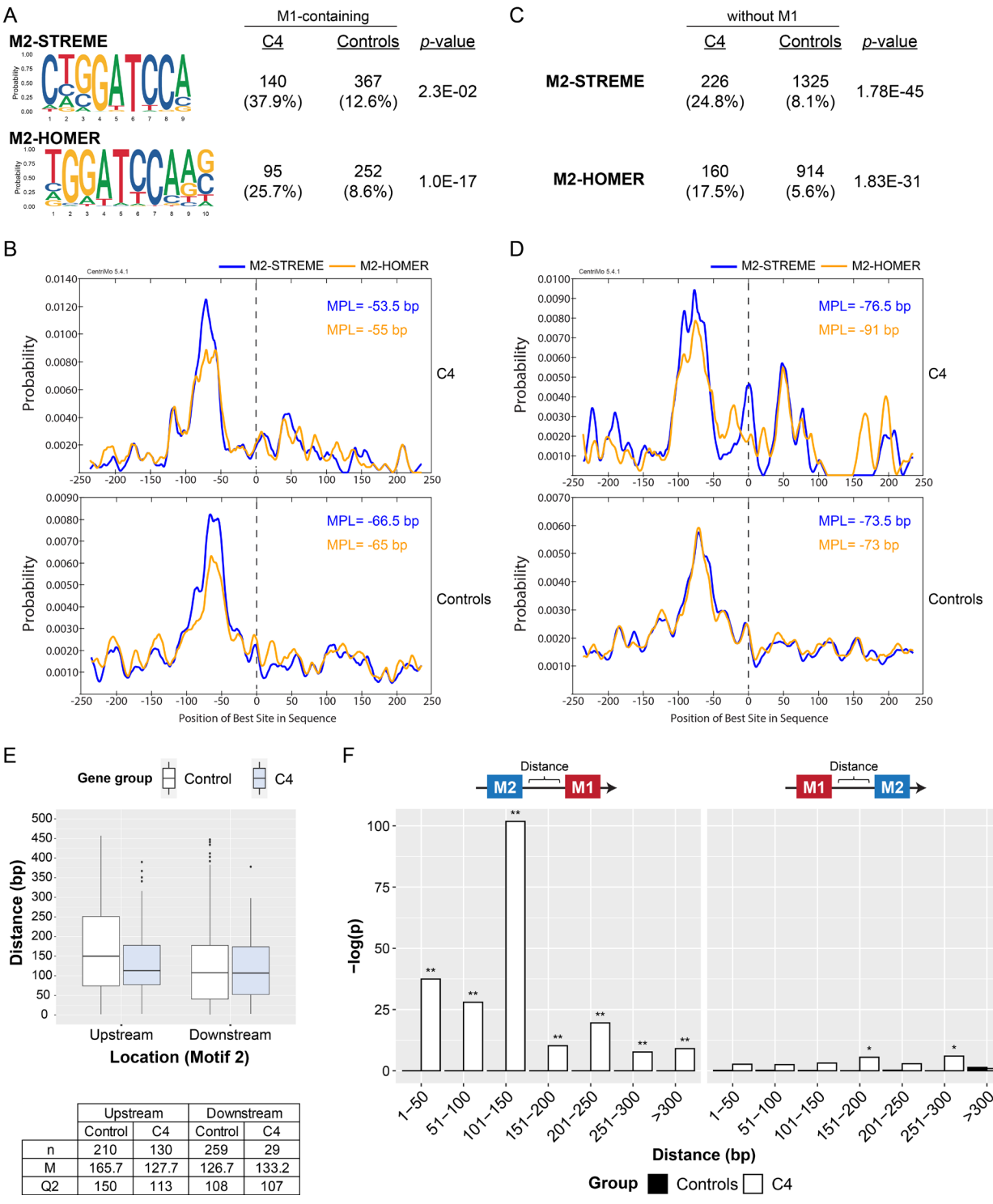


Figure 3.3. Characterization of motifs STREME-1 and HOMER-2 in Cluster 4 of *cht7*-DEGs.

(A) Identification of STREME-1 and HOMER-2 in Cluster 4 of *cht7*-DEGs (labeled C4). Enrichment of the STREME-1 and HOMER-2 motifs were computed by STREME and HOMER, respectively. Numbers and percentages of genes detected to contain STREME-1 or HOMER-2 in C4 and the control group are listed. See Table 3.2 and 3.3 for details. ***(B)*** Comparison between gene sets containing STREME-1 and HOMER-2 in C4 and the control group. ***(C)*** Gene examples showing the presence of STREME-1 (dark blue) and HOMER-2 (orange) and their distances to the transcription start site (TSS, green). ***(D)*** Positional distribution

Figure 3.3 (cont'd)

of STREME-1 (dark blue) and HOMER-2 (orange) in the TSS +/- 250 bp region of detected motif-containing genes of C4 and the control group. Y-axis: Probability of the motif occurrence derived from gene counts. X-axis: position of the TSS +/- 250 bp region. MPL: Maximum probability location.



	Upstream		Downstream	
	Control	C4	Control	C4
n	210	130	259	29
M	165.7	127.7	126.7	133.2
Q2	150	113	108	107

Figure 3.4. Enrichment of Motif 2 in the Motif 1-containing genes.

(A) Identification of Motif 2 (M2)-STREME and M2-HOMER in TSS +/- 250 bp regions of Motif 1 (M1)-containing genes of C4 by STREME and HOMER, respectively. Numbers and

Figure 3.4 (cont'd)

percentages of genes found to carry each motif in M1-containing genes of C4 and the control group are listed. ***(B)*** Positional distribution of M2-STREME (dark blue) and M1-HOMER (orange) along with the TSS +/- 250 bp region of genes carrying both M1 and M2 in C4 and the control gene sets. ***(C)*** Enrichment analysis of M2-STREME and M2-HOMER in TSS +/- 250 bp regions of C4 genes without the presence of M1. ***(D)*** Positional distribution of M2-STREME (dark blue) and M1-HOMER (orange) along with the TSS +/- 250 bp region of genes carrying only M2 in C4 and the control gene sets. ***(E)*** Distance analysis of M1 and M2 in the control group (white) and C4 gene set (light blue). X-axis: relative location of M2 to M1. Upstream: M2-M1. Downstream: M1-M2. The number of genes (n), average distance (μ), and median (Q2) of each gene set are labeled. ***(F)*** Statistical significance of the positional abundance of M1 and M2. The positional relationship between M1 and M2 of each panel is shown. X-axis: absolute distance between M1 and M2. Y-axis: -log(*p*-value) calculated from the hypergeometric test of each gene set. Refer to Materials and Methods for details. **: $p < 0.001$. *: $p < 0.01$.

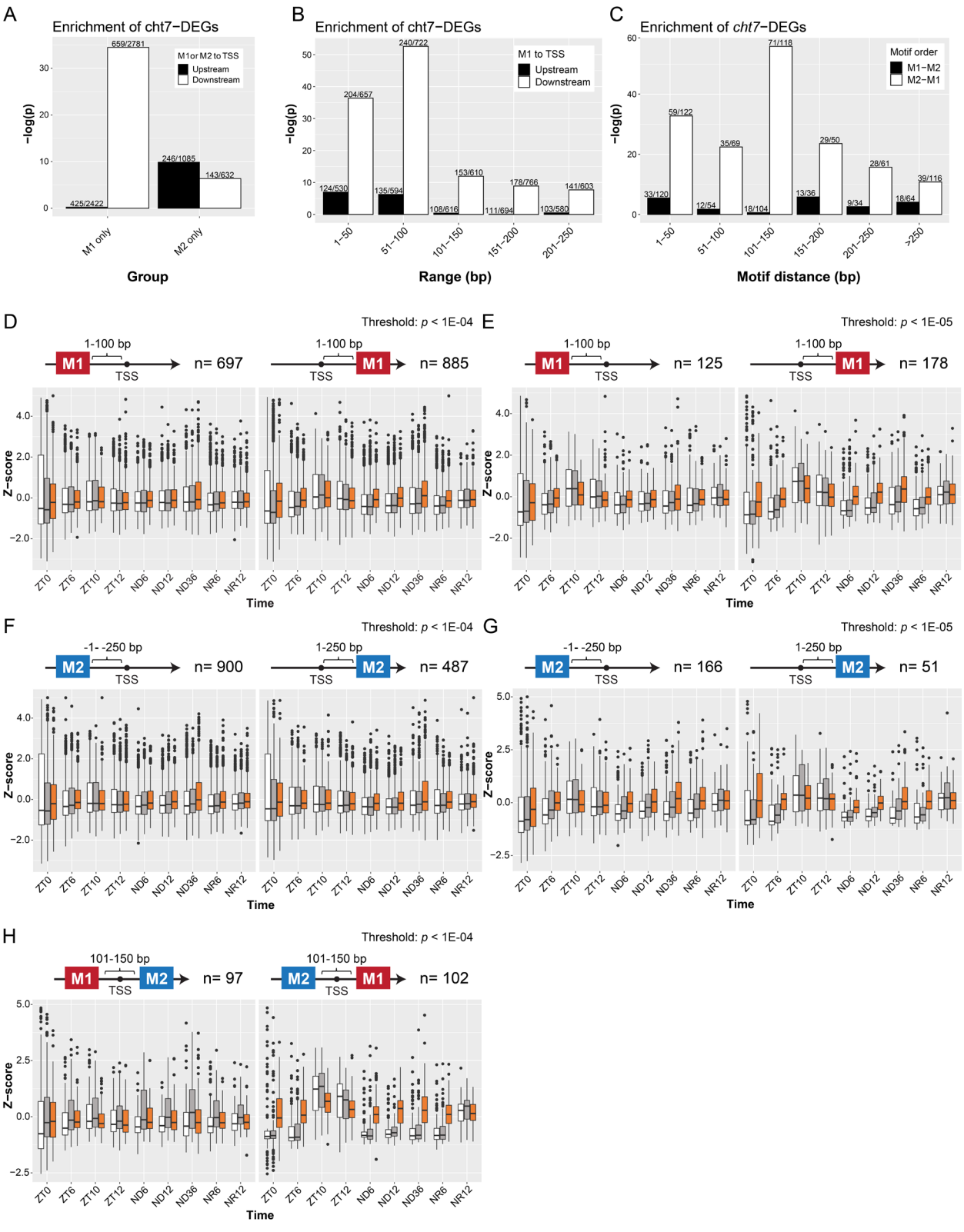
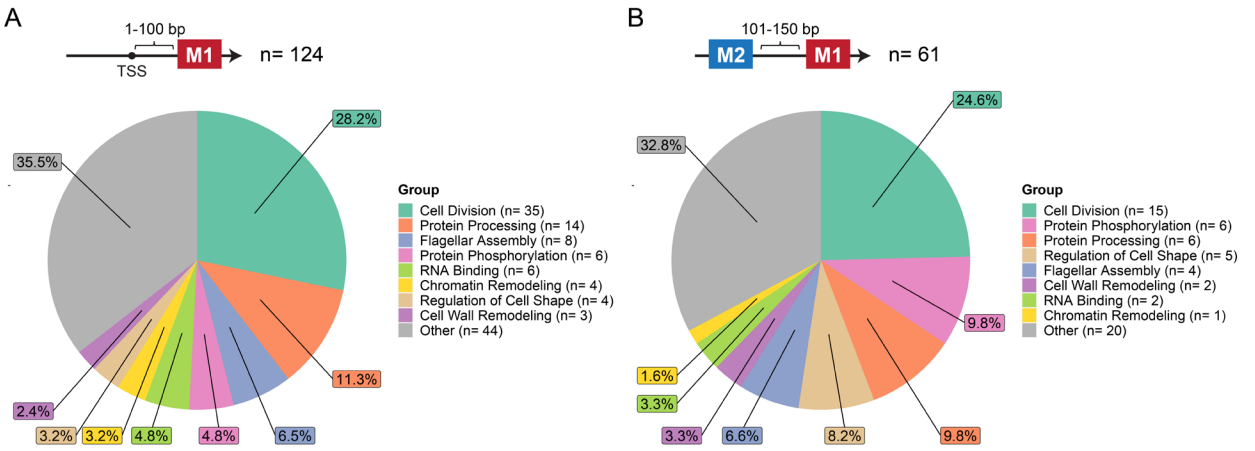


Figure 3.5. Genome-wide search and characterization of Motif 1 and 2-containing genes.

Figure 3.5 (cont'd)

(A) Enrichment of *cht7*-DEGs in genes containing only M1 and M2. The relative locations of each motif to TSS are colored as black (upstream of TSS) and white (downstream of TSS). ***(B)*** Enrichment of *cht7*-DEGs in M1 genes in accordance with the M1 positional distribution. The relative locations of M1 to TSS are colored black (upstream of TSS) and white (downstream of TSS). X-axis: absolute distance to TSS. ***(C)*** Enrichment of *cht7*-DEGs in genes of different M1 and M2 combinations. X-axis: absolute distance between M1 and M2. Y-axis for ***(A)*** to ***(C)***: The statistical significance is expressed as -log(*p*-value). ***(D)-(H)*** Boxplots showing the expression of genes with M1, M2, or both in WT (white), *CHT7HA::cht7* (grey), and *cht7* (orange). The positional relationships between M1, M2, and TSS are shown. Threshold *p*: *p*-value to determine if a sequence given is a specific motif. X-axis: Timepoints of N+, ND, and NR conditions. Y-axis: *Z*-score of gene expression.



C

Strand	Start	Sequence	End	Gene	Functional Category
+	58	GTCTTTGACTCG**AGCTGTC**ATAGGTGTCTTC	87	Cre01.g027400	Cell Division
-	74	CTGTCTTGTAGC**AGCTGTC**AGCCTAGCACGA	45	PSO2	Cell Division
-	86	TCTGCTGTCCCG**AGCTGTC**GGGAAAAGCAAG	57	SPO11A	Cell Division
-	75	GTGTGGGCCAGG**AGCTGTC**GAGCCTGGCTGC	46	KIN15-1	Cell Division
-	108	AGGTATGGCAGA**AGCTGTC**GAGTCCAGGTCA	79	SFI1-related	Cell Division
+	40	TGCGCTGTTTCC**AGCTGTC**TTAATCCTCCAC	69	Cre08.g374850	Cell Division
-	85	CGAGCTGGTGCA**AGCTGTC**AACAAAGCCTCG	56	CDC20	Cell Division
-	71	AGTGTACTCCTC**AGCTGTC**AACTCCTCCGGC	42	WEX	Cell Division
-	92	GGTCAAGTCACA**AGCTGTC**ACTGCAGCAAAT	63	KIN14A-2	Cell Division
-	96	ACCCTTCTTGCG**AGCTGTC**AGCCCTGGTCAG	67	SPC25-related	Cell Division
-	33	GAACATGATGGG**AGCTGTC**GTTGTTGCTAAT	4	ELG39	Protein Processing
-	84	CCCCGCTATTGC**AGCTGTC**AACAGAAAGATC	55	UBQ3	Protein Processing
-	65	GTCCAGTCTGCA**AGCTGTC**ACGCGTGCACAA	36	UBC24	Protein Processing
-	63	TCCAACTTCATG**AGCTGTC**GCTCCGTGGCGC	34	POC5	Flagellar Assembly
+	-3	TCACCTGCGCGA**AGCTGTC**GCTGGGCATAGC	26	CEP290	Flagellar Assembly
+	22	AATTTCTGTGTC**AGCTGTC**AACATTCGCCGG	51	SAS4	Flagellar Assembly
+	30	ATGGAACCACGG**AGCTGTC**ACTGTCGTTGAC	59	VFL1	Flagellar Assembly
+	12	AACATAGCTAGC**AGCTGTC**AGCTGTCTCAAG	41	SRR7	Cell Wall Remodeling
-	37	GCCAGTTCGCCG**AGCTGTC**GTCGTGACTGTT	8	CALS1	Cell Wall Remodeling
-	56	GGGCAATTCTCA**AGCTGTC**GCCAGTTCGCCG	27	CALS1	Cell Wall Remodeling
-	85	AGCGCGGGTGGC**AGCTGTC**GACGTCTGCGTG	56	MMP32	Cell Wall Remodeling
+	78	GAACCCTGCCCC**GGCTGTC**ACCTATTGTTTG	107	HTV1	Chromatin Remodeling
-	54	AGAGCTGACCCG**AGCTGTC**GGTCGCGCGCTT	25	HAV1	Chromatin Remodeling

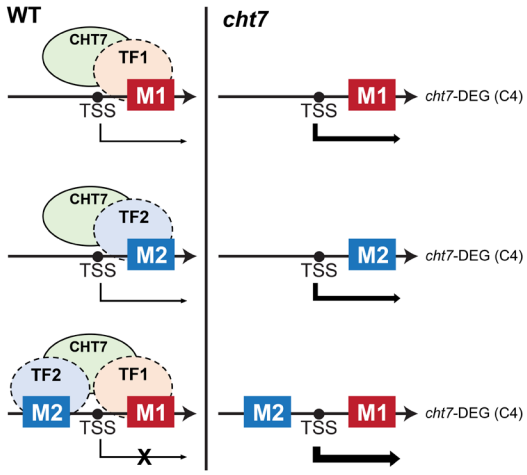


Figure 3.6. Biological functions of Motif 1 and Motif 2-Motif 1 containing genes.

(A) and ***(B)*** Pie charts showing annotated functions of genes with M1 in the TSS +1 to +100 bp region and M2-M1 with a 101 to 105 bp gap. ***(C)*** Representative examples of Motif 1 found in *cht7*-DEGs participating in cell division, protein processing, flagellar assembly, cell wall remodeling, and chromatin remodeling. The start and end sites indicate the relative location to TSS. ***(D)*** A working model illustrating the interaction between CHT7 and proposed TFs binding to Motif 1 and 2 on *cht7*-DEGs.

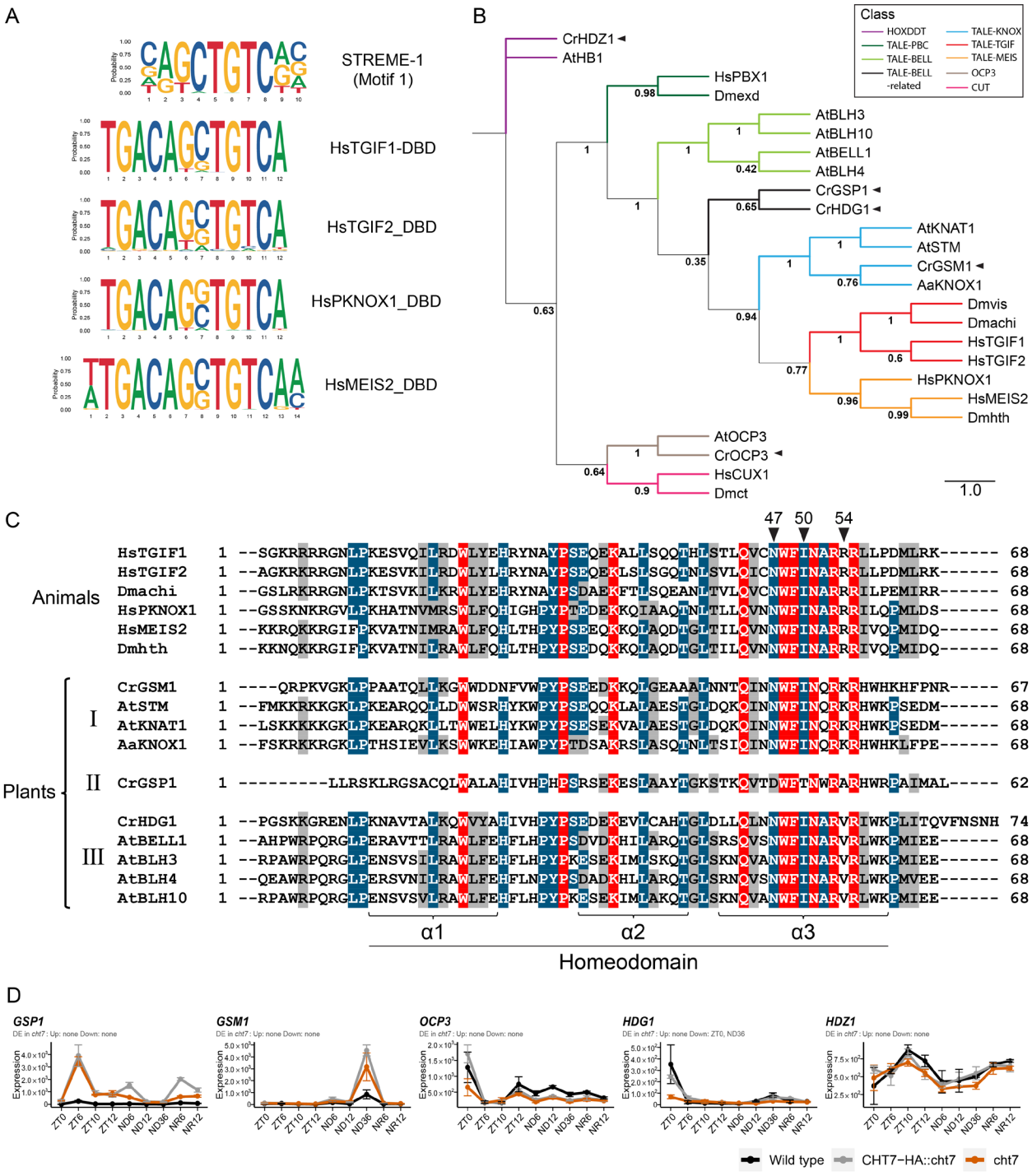


Figure 3.7. Comparison of homeoproteins between Chlamydomonas and other species.

(A) Sequence similarity between Motif 1 and TALE-HD targeted CREs. Data obtained from Tomtom. ***(B)*** Maximum likelihood phylogeny of homeoproteins of Chlamydomonas (n = 5; black triangle) and other species (n= 20). The analysis was done using PhyML. Clades of sub-classes

Figure 3.7 (cont'd)

are color labeled and shown in the legend. Confidence of branches was tested by approximate Likelihood-Ratio (aLRT) branch support. Cr: Chlamydomonas reinhardtii; At: Arabidopsis thaliana; Aa: Acetabularia acetabulum; Dm: Drosophila melanogaster; Hs: Homo sapiens. ***(C)*** Multiple sequence alignment of TALE-HDs. The three α-helixes in the homeodomain and key residues at positions 47, 50, 54 are labeled. The grouping of plant HDs is based on a previous study (Lee et al., 2008). ***(D)*** Expression profiles of Chlamydomonas genes encoding homeoproteins in *cht7* (orange), WT (black), and *CHT7-HA::cht7* (gray) across all timepoints. Y-axis: normalized read count.

REFERENCES

REFERENCES

Alexa, A., and Rahnenfuhrer, J. (2020). topGO: Enrichment analysis for Gene Ontology. R package version 2.42.0.

Alexa, A., Rahnenfuhrer, J., and Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. Bioinformatics 22:1600-1607.

Alvarez, J.M., Brooks, M.D., Swift, J., and Coruzzi, G.M. (2021). Time-based systems biology approaches to capture and model dynamic gene regulatory networks. Annu Rev Plant Biol 72:105-131.

Alvarez, J.M., Schinke, A.L., Brooks, M.D., Pasquino, A., Leonelli, L., Varala, K., Safi, A., Krouk, G., Krapp, A., and Coruzzi, G.M. (2020). Transient genome-wide interactions of the master transcription factor NLP7 initiate a rapid nitrogen-response cascade. Nat Commun 11:1157.

Anisimova, M., and Gascuel, O. (2006). Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 55:539-552.

Bailey, T.L. (2021). STREME: Accurate and versatile sequence motif discovery. Bioinformatics.

Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME Suite. Nucleic Acids Res 43:W39-49.

Belles-Boix, E., Hamant, O., Witiak, S.M., Morin, H., Traas, J., and Pautot, V. (2006). KNAT6: an Arabidopsis homeobox gene involved in meristem activity and organ separation. Plant Cell 18:1900-1907.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Pena-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133:1266-1276.

Bertolino, E., Reimund, B., Wildt-Perinic, D., and Clerc, R.G. (1995). A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. J Biol Chem 270:31178-31188.

Bodenhofer, U., Bonatesta, E., Horejs-Kainrath, C., and Hochreiter, S. (2015). msa: an R package for multiple sequence alignment. Bioinformatics 31:3997-3999.

Burglin, T.R. (1997). Analysis of TALE superclass homeobox genes (MEIS, PBC, KNOX, Iroquois, TGIF) reveals a novel domain conserved between plants and animals. Nucleic Acids Res 25:4173-4180.

Burglin, T.R., and Affolter, M. (2016). Homeodomain proteins: an update. Chromosoma 125:497-521.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 4:1184-1191.

Fischer, M., and DeCaprio, J.A. (2015). Does Arabidopsis thaliana DREAM of cell cycle control? EMBO J 34:1987-1989.

Geerts, D., Revet, I., Jorritsma, G., Schilderink, N., and Versteeg, R. (2005). MEIS homeobox genes in neuroblastoma. Cancer Lett 228:43-50.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59:307-321.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576-589.

Holland, P.W. (2013). Evolution of homeobox genes. Wiley Interdiscip Rev Dev Biol 2:31-45.

Imoto, I., Sonoda, I., Yuki, Y., and Inazawa, J. (2001). Identification and characterization of human PKNOX2, a novel homeobox-containing gene. Biochem Biophys Res Commun 287:270-276.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell 152:327-339.

Joo, S., Nishimura, Y., Cronmiller, E., Hong, R.H., Kariyawasam, T., Wang, M.H., Shao, N.C., El Akkad, S.E., Suzuki, T., Higashiyama, T., et al. (2017). Gene regulatory networks for the haploid-to-diploid transition of *Chlamydomonas reinhardtii*. Plant Physiol 175:314-332.

Kanrar, S., Onguka, O., and Smith, H.M. (2006). Arabidopsis inflorescence architecture requires the activities of KNOX-BELL homeodomain heterodimers. Planta 224:1163-1173.

Kurvari, V., Grishin, N.V., and Snell, W.J. (1998). A gamete-specific, sex-limited homeodomain protein in Chlamydomonas. J Cell Biol 143:1971-1980.

Laughon, A. (1991). DNA binding specificity of homeodomains. Biochemistry 30:11357-11367.

Lee, J.H., Lin, H., Joo, S., and Goodenough, U. (2008). Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. Cell 133:829-840.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature 276:565-570.

Lin, H., and Goodenough, U.W. (2007). Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, MID and MTD1. Genetics 176:913-925.

Lin, Y.T., Takeuchi, T., Youk, B., Umen, J., Sears, B.B., and Benning, C. (2022). Chlamydomonas CHT7 is involved in repressing DNA replication and mitotic genes during synchronous growth. G3 Genes|Genomes|Genetics, 12.

Marceau, A.H., Felthousen, J.G., Goetsch, P.D., Iness, A.N., Lee, H.W., Tripathi, S.M., Strome, S., Litovchick, L., and Rubin, S.M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun 7:12301.

Melhuish, T.A., Gallo, C.M., and Wotton, D. (2001). TGIF2 interacts with histone deacetylase 1 and represses transcription. J Biol Chem 276:32109-32114.

Mukherjee, K., Brocchieri, L., and Burglin, T.R. (2009). A comprehensive classification and evolutionary analysis of plant homeobox genes. Mol Biol Evol 26:2775-2794.

Muller, G.A., Wintsche, A., Stangner, K., Prohaska, S.J., Stadler, P.F., and Engeland, K. (2014). The CHR site: definition and genome-wide identification of a cell cycle transcriptional element. Nucleic Acids Res 42:10331-10350.

Reiser, L., Modrusan, Z., Margossian, L., Samach, A., Ohad, N., Haughn, G.W., and Fischer, R.L. (1995). The BELL1 gene encodes a homeodomain protein involved in pattern formation in the Arabidopsis ovule primordium. Cell 83:735-742.

Sadasivam, S., and DeCaprio, J.A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. Nat Rev Cancer 13:585-595.

Sadasivam, S., Duan, S., and DeCaprio, J.A. (2012). The MuvB complex sequentially recruits B-Myb and FoxM1 to promote mitotic gene expression. Genes Dev 26:474-489.

Scofield, S., Dewitte, W., and Murray, J.A. (2014). STM sustains stem cell function in the Arabidopsis shoot apical meristem and controls KNOX gene expression independently of the transcriptional repressor AS1. Plant Signal Behav 9.

Song, N., Zhu, X., Shi, L., An, J., Wu, Y., and Sang, J. (2009). Identification and functional analysis of a CDE/CHR element in the POLD1 promoter. Sci China C Life Sci 52:551-559.

Sudarsanam, P., Pilpel, Y., and Church, G.M. (2002). Genome-wide co-occurrence of promoter elements reveals a cis-regulatory cassette of rRNA transcription motifs in Saccharomyces cerevisiae. Genome Res 12:1723-1731.

Takeuchi, T., Lin, Y.T., Fekaris, N., Umen, J., Sears, B.B., and Benning, C. (2020a). Modulation of CHT7 complexes during light/dark- and nitrogen-mediated life cycle transitions of Chlamydomonas. Plant Physiol 184:1762-1774.

Takeuchi, T., Sears, B.B., Lindeboom, C., Lin, Y.T., Fekaris, N., Zienkiewicz, K., Zienkiewicz, A., Poliner, E., and Benning, C. (2020b). Chlamydomonas CHT7 is required for an effective quiescent state by regulating nutrient-responsive cell cycle gene expression. Plant Cell 32:1240-1269.

Tsai, C.H., Uygun, S., Roston, R., Shiu, S.H., and Benning, C. (2018). Recovery from N deprivation is a transcriptionally and functionally distinct state in Chlamydomonas. Plant Physiol 176:2007-2023.

Tsai, C.H., Warakanont, J., Takeuchi, T., Sears, B.B., Moellering, E.R., and Benning, C. (2014). The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in Chlamydomonas. Proc Natl Acad Sci U S A 111:15833-15838.

Whelan, S., and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18:691-699.

Wotton, D., and Taniguchi, K. (2018). Functions of TGIF homeodomain proteins and their roles in normal brain development and holoprosencephaly. Am J Med Genet C Semin Med Genet 178:128-139.

Yang, Y., Hwang, C.K., D'Souza, U.M., Lee, S.H., Junn, E., and Mouradian, M.M. (2000). Three-amino acid extension loop homeodomain proteins Meis2 and TGIF differentially regulate transcription. J Biol Chem 275:20734-20741.

Zhang, X., Li, J., Liu, A., Zou, J., Zhou, X., Xiang, J., Rerksiri, W., Peng, Y., Xiong, X., and Chen, X. (2012). Expression profile in rice panicle: insights into heat response mechanism at reproductive stage. PLoS One 7:e49652.

CHAPTER 4
Exploring the relationship of CHT7 to other cell cycle regulators

ABSTRACT

Previously, we have established a role for the CHC domain-containing protein CHT7 in the transcriptional regulation of the cell cycle in Chlamydomonas. In this chapter, I summarize our ongoing CHT7 studies exploring the potential relationship of CHT7 to other established cell-cycle regulatory mechanisms. We constructed double mutants lacking CHT7 and two components of the Retinoblastoma (RB/MAT3) protein complex. We then analyzed the potential effects of these mutants on cell cycle progression under synchronous growth. Results showed that although the CHT7 and RB complexes share some common gene targets, they might function independently to regulate gene expression. Furthermore, to advance our understanding of CHC proteins in the cell cycle regulation of Chlamydomonas, we characterized the mutant lacking a CHT7 paralog, which was tentatively named *chc-related 1 (chr1)* as well as the *chr1 cht7* double mutant. Preliminary experiments revealed that unlike CHT7, CHR1 negatively affects the transcription of genes involved in cell size control during the S/M phase, providing new insight into CHC-domain protein mediated cell cycle regulation.

INTRODUCTION

The extensive transcriptome analysis of the *cht7* mutant by our group has provided insights to the CHT7-mediated regulatory network of the cell cycle (Lin et al., 2022; Takeuchi et al., 2020b). For instance, CHT7 negatively affects the transcription of master cell cycle regulators such as cyclin B1 (CYCB1) and cyclin-dependent kinases CDKA1 and CDKB1 (Atkins and Cross, 2018; Lin et al., 2022; Tulin and Cross, 2015). On the other hand, the expression of genes encoding core proteins of the RB complex (RB, E2F, DP) was not influenced by the absence of CHT7 (Lin et al., 2022). Accordingly, we asked whether the RB-governed pathway overlaps with the gene network affected by the CHT7 complex. We were also curious to know whether the two protein complexes physically interact at specific points of the cell cycle.

In Chlamydomonas, the ortholog of RB, MAT3, is found in a complex composed of transcription factors DP1 and E2F. The active form of the MAT3 protein inhibits the activity of DP and E2F to activate their target genes, which in turn reduces cell division activity. The phosphorylation of MAT3 by cyclin D-dependent protein kinase CDKG1 removes the inhibitory effect of MAT3 on DP1 and E2F (Fang et al., 2006; Li et al., 2016; Umen and Goodenough, 2001). This regulatory mechanism is tightly associated with the cell size, as it only allows cells passing the cell-size checkpoint to enter the S/M phase and enables larger mother cells to divide more times than smaller ones during cell division. Hence, this process ensures daughter cells of a uniform size. This model is supported by the finding that the loss-of-function mutant of MAT3, *mat3*, exhibits a smaller cell size (median: 27 ± 2.4 μm^3) compared to wild type (median: 63 ± 3.1 μm^3) since the cell passes the cell-size checkpoint (commitment) prematurely and undergoes more-than-usual rounds of cell division during the S/M phase. On the other hand, the *dp1* mutant shows

a larger cell size (median: 104 ± 4.9 μm^3) due to the reduced rounds of cell division during the S/M phase (Fang et al., 2006).

As mentioned in previous chapters, CHT7 belongs to the CHC (CXC-Hinge-CXC) domain-containing protein family, some of which have been shown to regulate the cell cycle in animals (Marceau et al., 2016; Matsuo et al., 2012; Tsai et al., 2014), where interaction between the CHC and RB protein complexes has been reported. The animal CHC homolog LIN54 is a subunit of the nuclear-localized MuvB core complex (MuvB: Multi-Vulval Class B,) which alternatively interacts with other complexes, including the RB complex, to modulate gene expression for cell cycle progression (see review in Chapter 1) (Marceau et al., 2016; Sadasivam and DeCaprio, 2013; Sadasivam et al., 2012). Therefore, it is reasonable to assume that as a CHC protein, CHT7 might interact with the RB ortholog MAT3 in Chlamydomonas to regulate the cell cycle.

Our group conducted a series of protein assays to test the potential interaction between CHT7 and RB complexes. Immunoblots with blue native-PAGE showed that CHT7 and RB belonged to different protein complexes under mixotrophic conditions (Takeuchi et al., 2020a). Coimmunoprecipitation (Co-IP) using synchronized *HA-MAT3 CHT7-GFP::mat3-4 cht7* cells showed that HA-tagged MAT3 coimmunoprecipitated with CHT7-GFP when using a GFP-antibody during both G1 and S/M phase. However, the Co-IP signals were much weaker when using N-deprived (ND) cell cultures or performing a reverse Co-IP immunoprecipitating HA-MAT3 to detect CHT7-GFP in synchronized cells (Takeuchi et al., 2020a). Hence, the interaction between CHT7 and MAT3 appears to be unstable and difficult to detect under certain conditions. Furthermore, the negative control used in the Co-IP experiments was *HA-MAT3::mat3-4* for pull-down assays using the anti-GFP antibody or *CHT7-GFP::cht7* using the anti-HA antibody instead

of strains containing only the GFP or HA tag. Therefore, we cannot rule out the possibility that the detected interaction was between CHT7 and HA or MAT3 and GFP.

The discovery of CHT7 and follow-up studies highlighted the function of CHC proteins in the cell cycle regulation of Chlamydomonas. In addition to CHT7, there are two CHC paralogs which we tentatively designated as CHR1 (CHC-related 1, Cre12.g550250) and CHR2 (CHC-related 2, Cre08.g361400) in this study. As shown in Figure 1A (from Takeuchi et al., 2020a), multiple sequence alignment of CHC domains from Chlamydomonas and other species showed that all three CHC proteins in Chlamydomonas have a highly conserved CHC domain compared to other CHC proteins. The conserved features include the two cysteine-rich CXC motifs and two CXC-adjacent tyrosine (Y) residues that are essential for the DNA binding specificity in the animal CHC homolog LIN54 (Fig. 4.1A) (see Chapter 3 for more review) (Marceau et al., 2016; Takeuchi et al., 2020a). The expression of genes encoding CHR1 and CHR2 have been shown to be induced in the *cht7* mutant under both ND and NR conditions (Takeuchi et al., 2020b). Fig. 4.1B compares the transcriptional profiles of genes encoding CHT7, CHR1 and CHR2 in the *cht7* mutant and two control strains using data from a recent transcriptome analysis of cells grown under N-replete (N+), ND, and N-resupplied (NR) synchronous conditions (Lin et al., 2022). The first panel shows the increased expression of *CHT7* when cells are actively dividing. The second panel shows a similar timing of expression of *CHR1*, with slightly decreased expression in the *cht7* mutant at timepoints representing cell division. Notably, the *CHR1* gene is induced in *cht7* at non-cell dividing timepoints, including ZT6, ND6-36, and NR12. These data suggest an alternation of the transcription of CHR1 in the *cht7* mutant during different cell cycle stages. In contrast, the *CHR2* gene did not exhibit as significant a fluctuation in its expression pattern during the cell cycle nor was the expression altered in the *cht7* mutant (panel 3).

In this chapter, I focus on two research questions that arose from the above-mentioned findings: 1) Is there a relationship between the gene networks governed by the CHT7 and MAT3 (RB) complexes? If so, how does such interaction affect the cell cycle. 2) What is the function of CHR1 with regard to cell cycle progression and does CHR1 play a part in the CHT7-mediated pathway? To address the first question, we generated double mutants of CHT7 and MAT3 or DP1, namely *cht7 mat3-4* and *cht7 dp1*, to analyze their impacts on the progression of the cell cycle in algal cultures during synchronous growth. With a similar approach, we obtained the loss-of-function mutant of CHR1, *chr1*, and constructed the *chr1 cht7* double mutant to investigate the potential role of CHR1 in regulation of the Chlamydomonas cell cycle.

RESULTS

Construction and characterization of double mutants *cht7 mat3-4* and *cht7 dp1*

To investigate the potential interaction between CHT7 and the RB complex, we constructed two double-mutant strains, *cht7 mat3-4 (rb)* and *cht7 dp1* in the 21gr background (cw^+, mt^+). The resulting mutant strains were verified with PCR-based genotyping (Fig. 4.2A) and RT-qPCR probing of DNA sequences encoding CHT7, DP1, and RB in *cht7 dp1* (Fig. 4.2B) and *cht7 mat3-4* (Fig. 4.2C). As reported previously, cells of *dp1* are larger than 21gr (wild-type) due to a lower cell division activity, while the *mat3-4* cells are smaller than those of 21gr because of the premature entry into the S/M phase and a higher cell division number (Fig. 4.2F and 4.2G) (Fang et al., 2006; Umen and Goodenough, 2001). The *cht7 dp1* cells, however, were not bigger than 21gr, indicating that the less active cell division in *dp1* might be reversed by the *cht7* mutation which promotes cell division (Fig. 4.2H)(Lin et al., 2022; Takeuchi et al., 2020a). Cells of *cht7 mat3-4* were overall the same size as its parental lines (PLs) *cht7* and *mat3-4* (Fig. 4.2I).

Next, we examined the expression of genes encoding cell cycle regulators CYCA1 and CYCB1 (Atkins and Cross, 2018) in the *cht7 dp1* double mutant grown under synchronized conditions (Fig. 4.3A). As expected, genes encoding CYCA1 and CYCB1 were derepressed in *cht7* compared to 21gr and *dp1* at all stages of the cell cycle that were examined (Fig. 4.3A). In the wild-type line 21gr, CYCA1 and CYCB1 genes are expressed at ZT10, after cell enlargement, in preparation for cell division. In contrast, both *CYCA1* and *CYCB1* genes were failed to be turned on at ZT10 when DP1 is absent, supporting an assumed role of DP1 in activating some S/M genes (Fig. 4.3A) (Olson et al., 2010). However, when the *cht7* mutant was introduced to *dp1*, the resulting double mutant, *cht7 dp1*, accumulated transcripts of CYCA1 and CYCB1 throughout the cell cycle (Fig. 4.3A). This finding is in agreement with the cell size observation in Figure. 4.2F and 4.2H, suggesting the reduced expression of some cell cycle genes in *dp1* is compensated by their enhanced expression in the *cht7* mutant (Fig. 4.3A).

We next tested the expression of another cell-cycle marker gene, *CDKB1*, in the *cht7 mat3-4* mutant during synchronous growth (Fig. 4.3B). Compared to 21gr, the expression of the *CDKB1* gene at ZT0 (G0) was derepressed in the *mat3-4* and *cht7* mutants. In the *cht7 mat3-4* double mutant, *CDKB1* expression appears to be additive at all time points, with expression during S/M (at ZT12) approaching that of wild-type. Again, these results indicate that the protein complexes of CHT7 and RB might additively target the same genes.

Defective cell cycle progression in *cht7 mat3-4* but not in *cht7 dp1*

While cell growth and division seemed normal in *cht7 dp1*, we noticed that the *cht7 mat3-4* double mutant grew slower than single mutants lacking CHT7 or RB (Fig. 4.4A). Microscopic images of cell morphology showed that *cht7 mat3-4* cells frequently formed aggregated cell clumps (also

known as palmelloid clusters; de Carpentier et al., 2019) when grown under synchronous conditions (Fig. 4.4B). Therefore, it seemed possible that *cht7 mat3-4* cells have difficulty hatching from the mother cell after cell division, leading to delayed cell cycle progression. To test this possibility, we treated the synchronized cell cultures with autolysin, a cell-wall degrading enzyme secreted by Chlamydomonas gamete cells (Jaenicke et al., 1987). As shown in Figures 4.4C to 4.4E, the treatment of autolysin helped break up the large clumps of the *cht7 mat3-4* lines, releasing most of the daughter cells. We next tracked the growth of the released *cht7 mat3-4* cells with microscopy, finding that although most cells had enlarged, only a portion of them underwent cell division at ZT 13 and ZT14, representative timepoints of the S/M phase (Fig 4.4F and 4.4G). The cells of *cht7 mat3-4* resumed forming palmelloids after a short period of recovery from the autolysin treatment (Fig. 4.4H).

We further investigated the cell cycle progression of *cht7 dp1* and *cht7 mat3-4* by measuring cell size changes throughout the ZT stages of growth in N-replete media. The result showed that during synchronous growth, *cht7 dp1* cells behaved similarly to the wild type 21gr, as they gained in size during the G1 phase and entered the S/M phase at ZT12-14 to produce daughter cells (Fig. 4.5A). This finding was further confirmed by a detailed measurement of mitotic activity by microscopy performed in the collaborating lab of James Umen, which showed that the *cht7 dp1* mutant had elevated mitotic activity at both ZT0 (G0) and ZT 12 (S/M) compared to 21gr (Fig. 4.5B). For *cht7 mat3-4* cells treated with autolysin, there was only a moderate increase in cell size compared to two other strains (Fig. 4.5A), and the timing of mitosis was significantly delayed (Fig. 4.5B). From these findings, we conclude that the progression of the cell cycle is not reduced in the *cht7 dp1* double mutant, which might even have shown enhanced mitotic activity. In contrast, in *cht7 mat3-4,* cell growth is impaired, although the actual cause is not clear.

Construction of the chr1 and chr1 cht7 mutant strains

As mentioned previously, there are two additional CHC proteins in Chlamydomonas other than CHT7. Our recent RNA-seq analysis showed that in control strains 6145c (WT mt^-) and *CHT7HA::cht7,* the gene encoding CHR1 is mainly expressed during the S/M phase, implying CHR1 might participate in S/M-related processes (Fig. 4.1B). Furthermore, the expression of *CHR1* was slightly induced in *cht7* compared to the two controls (Fig. 4.1B). These findings suggested that CHR1 is subject to the CHT7-regulated gene network or was induced as negative feedback in response to the loss of CHT7. To explore these possibilities, we obtained an insertional mutant strain of *CHR1*, *chr1*, from the CLiP mutant library (LMJ.RY0402_188955; Fig. 4.6A). We next backcrossed *chr1* with several wild-type strains to remove the background mutations and obtain a *chr1* strain with an intact cell wall (see Materials and Methods for details). We then introduced the *cht7* mutation by crossing into *chr1* to create the *chr1 cht7* double mutant. All strains were validated using PCR-based genotyping and RT-qPCR analysis (Fig. 4.6B and 4.6C).

Growth-related phenotypes of *chr1* and *chr1 cht7* during photoautotrophic growth

We first examined whether the loss of CHR1 affects cell growth in response to changes in nitrogen availability, because this is a hallmark of phenotypes associated with the *cht7* mutation. Cells of *chr1* and *chr1 cht7* (two lines per strain) were grown photoautotrophically under N+ and NR conditions following 48 h of N deprivation, alongside the PLs wild type 6145C and *cht7*. For the N+ *chr1* cell cultures, the two lines exhibited a higher cell density than other strains after growing for four days, and the cells at ZT0 were smaller than the wild type (Fig. 4.7A and 4.7B). As shown in Fig.4.8, microscopic images of the *chr1 cht7* double mutant showed that some cells were aggregated, interfering with the results of cell size and density measurements with the Coulter

counter (Fig. 4.7A). For NR cell cultures, the growth curves of *chr1* and WT were similar, suggesting that exit from N deprivation-induced quiescence to the normal cell cycle is not impaired in *chr1* in contrast to *cht7* (Fig. 4.7C). On the other hand, both *chr1 cht7* and *cht7* had a slow recovery of growth following N resupply (Fig. 4.7C). These observations suggested that CHR1 might not participate directly in the regulation of quiescence as CHT7 does.

The higher cell density and smaller cell size observed for N-replete cell cultures of *chr1* indicated that the progression of the cell division cycle might be affected by the absence of CHR1. To test this hypothesis, we grew *chr1*, *chr1 cht7*, and their PLs in ePBRs under synchronous conditions with a turbidity control, as described previously (Lin et al., 2022; Takeuchi et al., 2020a). Consistent with the previous finding, the ePBR-grown ZT0 cells of *chr1* were generally smaller than WT, although some cell clumps were observed in *chr1* line 1-12-1 (Fig. 4.7D). Together, we showed that CHR1 might affect the cell size by promoting the activity of cell division.

Analyses of the cell cycle progression in *chr1* and *chr1 cht7*

Figure 4.8 shows the cell cycle progression of ePBR cultures of *chr1* and *chr1 cht7* during N-replete synchronous growth. The wild-type cell entered the S/M phase at ZT12 and completed cell division within 2 to 4 hours, as reported previously (Fig. 4.8A) (Lin et al., 2022). The *chr1* and *cht7* mutants also exhibited mitotic activities at ZT12, and some cells showed a prolonged phase of cell division which lasted for at least 4 hours (Fig. 4.8B to 4.8D). Moreover, a delay in hatching was observed in some *chr1* cells, which did not occur until ZT4 to ZT8 the next day (Fig. 4.8C and 4.8D). For the *chr1 cht7* double mutant, the entry into S/M phase was delayed to ZT14, but most cells seemed to complete cell division within 2 to 4 hours. However, most *chr1 cht7* cells remained in a 4-cell cluster structure (palmelloid) at all sample times, and single cells were rarely

observed (Fig. 4.8E). The occurrence of palmelloids in Chlamydomonas as a stress response is considered in the discussion (de Carpentier et al., 2019).

To better understand the transcriptional responses of S/M-specific processes to the absence of CHR1, we performed RT-qPCR analyses of three representative sets of cell cycle genes using the same synchronized cell cultures. Proteins encoded by each gene set are known to participate in a specific process of the S/M phase. CYCD2 and CDKG1 have been reported to modulate the activity of cell division in accordance with the cell size (Li et al., 2016). POLA1 is a subunit of the DNA polymerase, which is directly involved in DNA replication, and CDKB1 plays an essential role in mitotic entry (Atkins and Cross, 2018; Tulin and Cross, 2015). APC1 and APC6 are two protein components of the Anaphase Promoting Complex (APC) which is crucial for chromosome segregation and mitotic exit (Atkins and Cross, 2018).

In WT, genes encoding CYCD2 and CDKG1 were expressed mainly at ZT10, while the peak expression of *CDKG1* was delayed to ZT12 in *chr1* and *chr1 cht7* (Fig. 4.9A). The expression of *CYCD2* was induced in *chr1* at ZT10 and ZT12, while in *chr1 cht7,* its expression peaked at ZT12 (Fig. 4.9A). Overall, mRNAs of CDKG1 and CYCD2 were more abundant in *chr1* than WT during the S/M phase, which might be an indication that they are responsible for the reduced cell size and prolonged stage of cell division observed in *chr1* (Fig. 4.7 and Fig. 4.8B to 4.8D). Furthermore, the delay in peak expression of genes encoding CYCD2 and CDKG1 in *chr1 cht7* might contribute to the late S/M entry of the double mutant (Fig. 4.8E).

As reported previously, genes encoding POLA1 and CDKB1 were misregulated in *cht7* at non-cell dividing timepoints such as ZT0, ZT4, and ZT6, which was not seen in *chr1* (Fig. 4.9B) (Lin et al., 2022). On the other hand, genes encoding APC1 and APC6 were abnormally induced in both *chr1* and *chr1 cht7* at ZT10 and ZT12 compared to WT and *cht7* (Fig. 4.9C). Together,

these results indicated that CHR1 might regulate a different set of S/M genes than CHT7 (Fig. 4.9D).

DISCUSSION

We inspected the relationships of CHT7 to other known or potential cell cycle regulators. First, we explored the interaction between RB- and CHT7-mediated regulatory networks of the cell cycle. Second, we conducted preliminary experiments to investigate the role of a CHC protein paralog of CHT7, CHR1, in programming the transition between cell cycle stages.

We generated two double mutant strains, *cht7 mat3-4* and *cht7 dp1* (Fig. 4.2). Microscopic observations showed that the large-cell phenotype of *dp1* appeared to be rescued by the absence of CHT7 (Fig. 4.2F to and 4.2H). The abnormal cell size in *dp1* is caused by the reduced transcriptional activation of cell division (Fang et al., 2006; Lin et al., 2022), while *cht7* has been shown to derepress genes for cell division. Therefore, it is reasonable to conclude that the de-repression of cell cycle genes in *cht7* compensates for the opposite effect of *dp1*, thus restoring normal S/M activity. This conclusion is supported by the finding that the down-regulation of genes encoding CYCA1 and CYCB1 in *dp1* was reversed in the *cht7 dp1* double mutant (Fig. 4.3A). These results also suggested that the CHT7 and DP-mediated gene regulations are two parallel processes targeting a similar gene set.

The transcriptional activation of *CDKB1* is crucial for the entry into mitosis during the S/M phase (Atkins and Cross, 2018). We found that *CDKB1* expression was abnormally induced in both *cht7* and *mat3-4* at ZT0 (G0 phase) compared to 21gr (wild type) and was further increased in *cht7 mat3-4* (Fig. 4.3B). Again, this finding suggests that RB and CHT7 complexes act independently to negatively regulate the transcriptional program during the S/M phase (Fig. 4.5C).

This logic is supported by the finding that the two proteins belong to different protein complexes while both negatively regulate genes for cell division (Fang et al., 2006; Takeuchi et al., 2020a). Even though RB and CHT7 interact at specific points of the cell cycle, there is no strong evidence supporting the idea that the RB-CHT7 interaction is required for the transition between cell cycle stages.

While we could not validate a coordination between RB and CHT7 in cell cycle regulation, the deletion of both genes did cause additional stress in the double mutant. The *cht7 mat3-4* strains were found to grow slower than the PLs (Fig. 4.4A). Furthermore, *cht7 mat3-4* cells grown under synchronous conditions tended to form cell clumps also known as palmelloids, whose presence is an indicator of cellular stress in Chlamydomonas (Fig. 4.4B and 4.4C) (de Carpentier et al., 2019). To examine whether the slow growth of *cht7 mat3-4* is caused by an artifact in the size exclusion of the cell density measurements, we treated cell cultures of synchronous growth with gamete autolysin, a cell-wall degrading enzyme secreted during gamete fusion (Jaenicke et al., 1987). The result showed that autolysin treatment was able to aid the release of daughter cells from the mother cell wall (Fig. 4.4D and 4.4E). However, only a portion of cells underwent cell division at ZT12 and ZT14 (S/M phase). Microscopic analysis also showed that entry into the S/M phase might be delayed in the *cht7 mat3-4* double mutant (Fig. 4.4F, 4.4G, and 4.5). In conclusion, these findings suggest that the impaired growth of *cht7 mat3-4* cells is not due to problems with hatching. Instead, it is likely that the abnormally induced S/M program in *cht7 mat3-4* caused some negative effects like replicative stress, which inhibited cell growth and led to palmelloid formation.

In addition to the relationship between CHT7- and RB-regulated gene networks, we investigated the potential regulatory role of a previously uncharacterized CHC protein, CHR1, in cell cycle progression. Gene expression analysis with RT-qPCR revealed that the expression of

genes encoding CDKG1 and CYCD2 were overall higher in the *chr1* mutant than in the wild type under synchronous conditions (Fig. 4.9C). The functions of D-type cyclins and CDKG1 in mitotic cell size control have been studied previously (Li et al., 2016). Interestingly, a transgenic strain overexpressing *CDKG1* also exhibited a smaller cell size than wild-type cells, which is consistent with the finding that cells of *chr1* appeared to be smaller than the wild type (Fig. 4.7B and 4.7D) (Li et al., 2016). These findings suggest that CHR1 negatively affects the transcription of genes for cell size control. Additionally, the peak gene expression of *CYCD2* and *CDKG1* was delayed by 2 hours in *chr1 cht7* compared to the wild type, which is consistent with the double mutant's delayed entry into the S/M phase (Fig. 4.7A and Fig. 4.8). In contrast, genes encoding POLA1 and CDKB1 were not significantly altered in *chr1* compared to *cht7*, while the transcript levels of *APC1* and *APC6* were induced in *chr1* but not *cht7* at ZT12 (Fig. 4.9B and 4.9C). Together, these results indicated that CHT7 and CHR1 might regulate different S/M processes, as depicted in Fig. 4.9D. The impact of the misregulated transcription programs in *chr1* and *chr1 cht7* on the cell cycle remains unclear. Nonetheless, the finding that both *chr1* and *chr1 cht7* all showed the stress-responsive palmelloid structures suggests that the deletion of CHR1 may induce stress responses to the cell (de Carpentier et al., 2019), which was further induced in the double mutant. It is possible that the loss-of-function of CHC proteins, a general repressor of cell cycle genes, could lead to replicative stress during non-cell dividing stages. Future work will involve CHR1 complementation analysis, subcellular localization of the CHR1 protein, and global transcriptome analysis of the *chr1* mutant. A detailed examination of the mitotic peak and division number would also provide further information on the involvement of CHR1 in cell division.

MATERIALS AND METHODS

Strain Construction and Validation

The construction of *cht7* (cw^+) was described previously (Takeuchi et al., 2020b). *cht7 dp1* was constructed by crossing *cht7* (cw^+ mt^-; line 1-10-1) with *dp1* (cw^+ mt^+; provided by James Umen). *cht7 mat3-4* was created by crossing *cht7* (cw^+ mt^-; line 2-4-3A) with *mat3-4* (cw^+ mt^+; provided by James Umen). The resulting progenies were verified with PCR-based genotyping using crude cell extracts with Q5® High-Fidelity 2X Master Mix (M0492S, New England BioLabs) and primers indicated in Table 4.1.

The *chr1* mutant (cw^+; strain id: LMJ.RY0402.188955) was ordered from the CLiP mutant library (https://www.chlamylibrary.org/index). The insertion of the *CIB1* cassette into exon5 of *CHR1* (Cre12.g550250) was verified with PCR-based genotyping as described above. After verification the original *chr1* strain (cw^+ mt^-) was further crossed with CC-112 (cw^+ mt^+) to remove a second insertion in Cre09.g391650. The resulting 2nd generation *chr1* (cw^+ mt^-; line 1-4-3) was further crossed with 21gr (cw^+ mt^+) to obtain the 3rd generation *chr1* (cw^+ mt^+; line 1-9-1). Subsequent crosses and backcrosses to 6145c (cw^+ mt^-) produced the 5th generation of *chr1* (cw^+ mt^-), while crosses with *cht7* (cw^+ mt^-; line 2-4-3A) produced progeny selected for *chr1 cht7* (cw^+ mt^-).

Real-Time Quantitative PCR (RT-qPCR)

Total RNA was extracted using the RNeasy Plant Mini Kit (74904) and QIAcube Connect (9002864) from Qiagen. DNase treatment and reverse transcription of 0.5-1 µg of extracted RNA were performed using the SuperScript IV VILO Master Mix (11756050, Thermo Fisher Scientific). RT-qPCR was performed using Fast SYBR™ Green Master Mix (4385612 Thermo Fisher

Scientific) in the Applied Biosystems™ 7500 Fast Real-Time PCR System (Thermo Fisher Scientific). Data analysis was done with R version 4.0.0.

Growth Conditions

For non-ePBR synchronous growth, cells were inoculated in 250 mL bottles containing 100 mL of N-replete high salt (HS+N) medium (Sueoka, 1960) under 12 h light: 12 h dark cycles with a light intensity of 150 μmol $m^{-2}s^{-1}$, and 5% CO2 was pumped into the cell culture regularly (30 seconds every 15 minutes). For N deprivation, the non-ePBR cell cultures during N-replete synchronous growth were pelleted at 3000 g for 3 to 5 minutes, rinsed and resuspended with N-deplete high salt (HS-N) medium to 1×10^6 cells/mL in a total of 100 mL, with cells cultured under the same growth conditions as above for 48 hours. For N resupply, filter-sterilized 1M NH_4Cl was added to ND cell cultures to reach a final concentration of 100 mM. Methods for ePBR synchronous growth are described in Chapter 2.

Autolysin Treatment

The autolysin solution was provided by James Umen. The entire ePBR culture (330 mL) at ZT-1 on Day 12 was collected and pelleted at 3000 xg for 5 minutes in the presence of 0.005% Tween-20. The cell pellet was resuspended with 10 mL of 1:1 dilution of autolysin with HS-N medium and incubated under dark for 30 minutes. Next, the cell pellet was washed with 50 mL of HS-N medium twice and resuspended with 330 mL of HS+N medium to be re-inoculated to ePBR.

Microscopy and Cell Size Analysis

Cell samples collected from the cell cultures above were fixed with 0.2 % glutaraldehyde at 4 ºC for at least 10 minutes prior to analysis. Cell density and cell-size distribution were obtained using a Beckman Coulter Counter Z2 (Beckman Coulter), and the resulting data were analyzed with the Z2 software included and R version 4.0.0. Microscopy analysis was done with the Leica Inverted Microscope Solution DMi8 S Platform (Leica Microsystems Inc, IL, USA) equipped with a 63x oil immersion objective (HC PL APO 63x/1.40 Oil CS2). Microscopy images were taken using the LAS software included in the Leica system.

ACKNOWLEDGEMENTS

We thank Dr. James Umen at Donald Danforth Plant Science Center for kindly providing the *mat3-4* and *dp1* strains, the autolysin solution, and for the cell division quantifications in Fig. 4.5B. The crosses and maintenance of strains described in Materials and Methods were done by Dr. Barbara B. Sears from the Department of Plant Biology, Michigan State University.

APPENDIX

Primer	Sequence	Note
RB-GT-F2	CAGAGCGCGGGCTTGTAGCACTAA	*RB/MAT3* genotyping
RB-GT-R2	AAAGAAGTCCAGCAGGCTGCGC	
APH7-F	CTCAAGTGCTGAAGCGGTAG	*cht7* insertion genotyping (Tsai et al., 2014)
APH7-R	TGGTGGTCGACAAACTCTTG	
CHT7-GT-F1	TCTCGGGTCAGATATTGGG	*CHT7* genotyping (Tsai et al., 2014)
CHT7-GT-R1	TTGAGGCAGAACGACTTCTTG	
MTP(+)-F	GCTGGCATTCCTGTATCCTTGACG	*Mating-type plus* genotyping (Takeuchi et al., 2020b)
MTP(+)-R	GCGGCGTAACATAAAGGAGGGTCG	
MTM(-)-F	CGACGACTTGGCATCGACAGGTG	*Mating-type minus* genotyping (Takeuchi et al., 2020b)
MTM(-)-R	CTCGGCCAGAACCTTTCATAGGGTGG	
DP1-GT-F1	GCCATACCGCCACTGCATT	*DP1* genotyping
DP1-GT-R1	GCATTTCACAGCGCTGCATCT	
Cre12.g550250-GT-F2	CTGGGCATCAACTTGGACTC	*CHR1* genotyping 188955-IB + R2: *chr1* insertion
Cre12.g550250-GT-R2	TATGCCTGCAGCTGATTCTG	
188955-IB	CGGGGGCCAGGCGGCTCATGCA	
RB1-qF1	TTTCTCGTGCTGCGAGTC	RB/MAT3 RT-qPCR
RB1-qR1	GGCAAGGATGGCGAAGAT	
CBLP-qF	GCCACACCGAGTGGGTGTCGTGCG	CBLP RT-qPCR (Tsai et al., 2014)
CBLP-qR	CCTTGCCGCCCGAGGCGCACAGCG	
CHT7-qF2	TGCAACTGCAAGAAGTCGTT	CHT7 RT-qPCR
CHT7-qR2	CGGCACTCCACACACTTG	
CYCA1-qF1	CTGGGTGACCTGACAAACAA	CYCA1 RT-qPCR (Takeuchi et al., 2020b)
CYCA1-qR1	GAATGGAAGGCTGGTGGAA	
CYCB1-qF1	AGCGACTACATGACCAAGCAGACC	CYCB1 RT-qPCR (Takeuchi et al., 2020b)
CYCB1-qR1	TTCAGGAAGCGGTCGATGAGGTTC	
POLA1-qF1	TCGGCGAGGATGACTACTT	POLA1 RT-qPCR
POLA1-qR1	CCTTGCCTTGCTCCTTCTT	
CDKB1-qF	GACAACGCTGCGTGAGATTTC	CDKB1 RT-qPCR (Fang et al., 2006)
CDKB1-qR	ACCAGGTAAAGGCATGGCTTG	
Cre12.g550250-q1F	AGGACAAAGGACGAGGACTA	CHR1 RT-qPCR (q1)
Cre12.g550250-q1R	TTTGCAGTTGCATCGCTTG	
Cre12.g550250-q2F	CCCAAACCCAACTCCATCA	CHR1 RT-qPCR (q2)
Cre12.g550250-q2R	CCTCCGCCACTAAGGAAAG	
CDKG1-qF	CGAGCTAGCGCCTCTTATTC	CDKG1 RT-qPCR
CDKG1-qR	CACTGTGGTTGTGAAGTTGTG	
CYCD2-qF	GGCAACAAGAAGCCACTGTA	CYCD2 RT-qPCR
CYCD2-qR	GTTGTGGTACTGACGCAGAAA	
APC1-qF	TGCCGCTGCTATCCTTTC	APC1 RT-qPCR
APC1-qR	CCGTCGCTCCGCTTAAA	
APC6-qF	GAGCTCCGGGAGATGTTTATC	APC6 RT-qPCR
APC6-qR	GGCCAGGTACACAAACTCA	

Table 4.1. Primers used for genotyping PCR and RT-qPCR

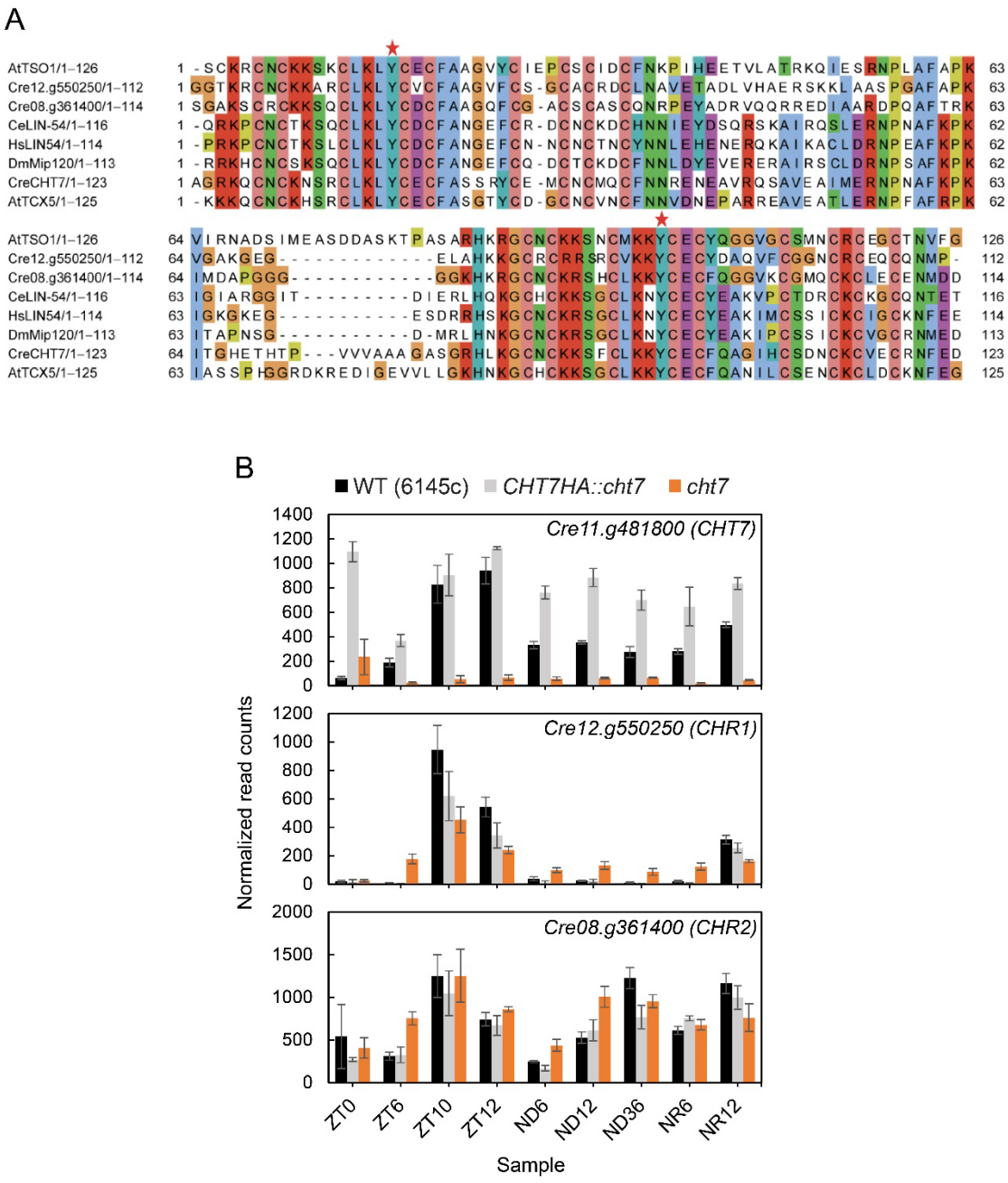


Figure 4.1. CHC proteins in Chlamydomonas.

(A) Multiple sequence analysis of the CHC domain in Chlamydomonas and other species (Takeuchi et al., 2020) **(B)** Transcript levels of CHC protein-encoding genes CHT7, CHR1 and CHR2 in the wild type 6145c (black), complemented line *CHT7HA::cht7* (grey), and *cht7*

Figure 4.1 (cont'd)

mutant (orange) during N-replete (N+) synchronous growth. *CHR1: CHC-related 1*; *CHR2: CHC-related 2*. Y-axis: Normalized read counts. ZT(n) samples represent cells under normal N+ conditions, where n = hours after the onset of light. ND(n) samples represent N-deprived (ND) cells, where n = hours after N removal. NR(n) samples represent N-resupplied ND cells, were n = hours of N resupply.

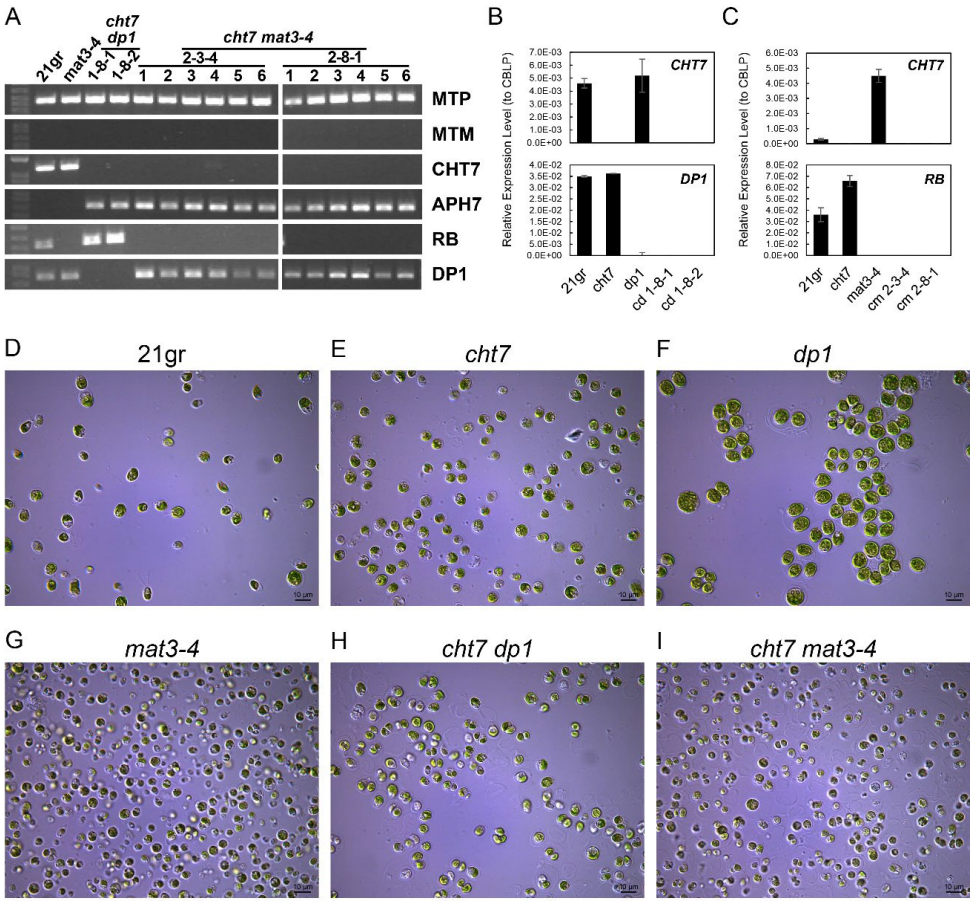


Figure 4.2. Characterization of double mutants *cht7 dp1* and *cht7 mat3-4*.

(A) PCR genotyping of the *cht7 dp1* and *cht7 mat3-4* double mutants and parental lines (PLs). The PLs include the wild type 21gr and the single mutant lines. Gene markers are listed below. Mating type: *MTP* (*Mating-type plus*) and *MTM* (*Mating-type minus*); *cht7*: *CHT7* and *APH7* (*Aminoglycoside Phosphotransferase Gene 7*, insertion marker); *mat3-4*: *RB*; *dp1*: *DP1*. Because of the complete linkage between mat3-4 and the mating-type locus, all strains are mt^+. ***(B-C)*** RT-qPCR analysis of ***(B)*** *CHT7* and *DP1* in *cht7 dp1* (*cd*) and ***(C)*** *CHT7* and *RB* in *cht7 mat3-4* (*cm*). PLs (21gr, *cht7, dp1*) are included. Each strain has two individual independently isolated lines. Cells were grown under synchronous conditions. ***(D-I)*** Microscopic images of cells of the PLs, *cht7 dp1*, *cht7 mat3-4*, and grown in TAP medium for 10 days. Scalebar = 10 µm.

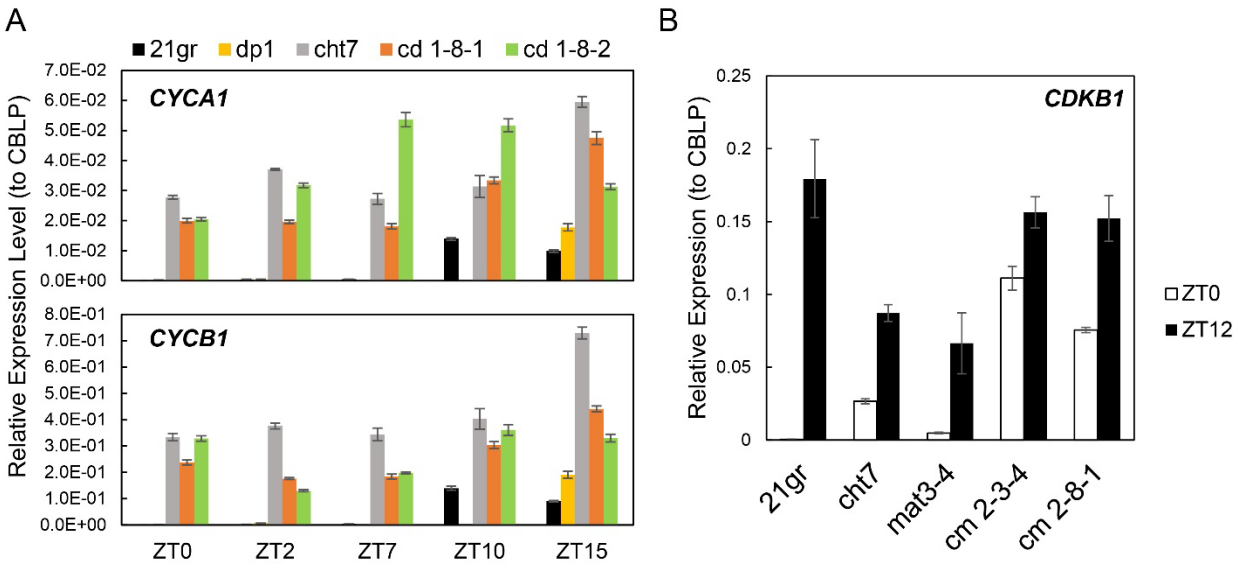


Figure 4.3. Gene expressions of cell cycle markers in synchronized *cht7 dp1* and *cht7 mat3-4*

(A) RT-qPCR analyses of CYCA1 and CYCB1-encoding genes in *cht7 dp1* (*cd*, 1-8-1: orange; 1-8-2: green) and PLs (21gr: black; *dp1*: yellow; *cht7*: grey) during N-replete synchronous growth. ***(B)*** RT-qPCR analysis of CDKB1-encoding gene in *cht7 mat3-4* (*cm*) and PLs during N-replete synchronous growth (ZT0: white; ZT12: black). ZT(n) samples represent cells under normal N+ conditions, where n = hours after the onset of light.

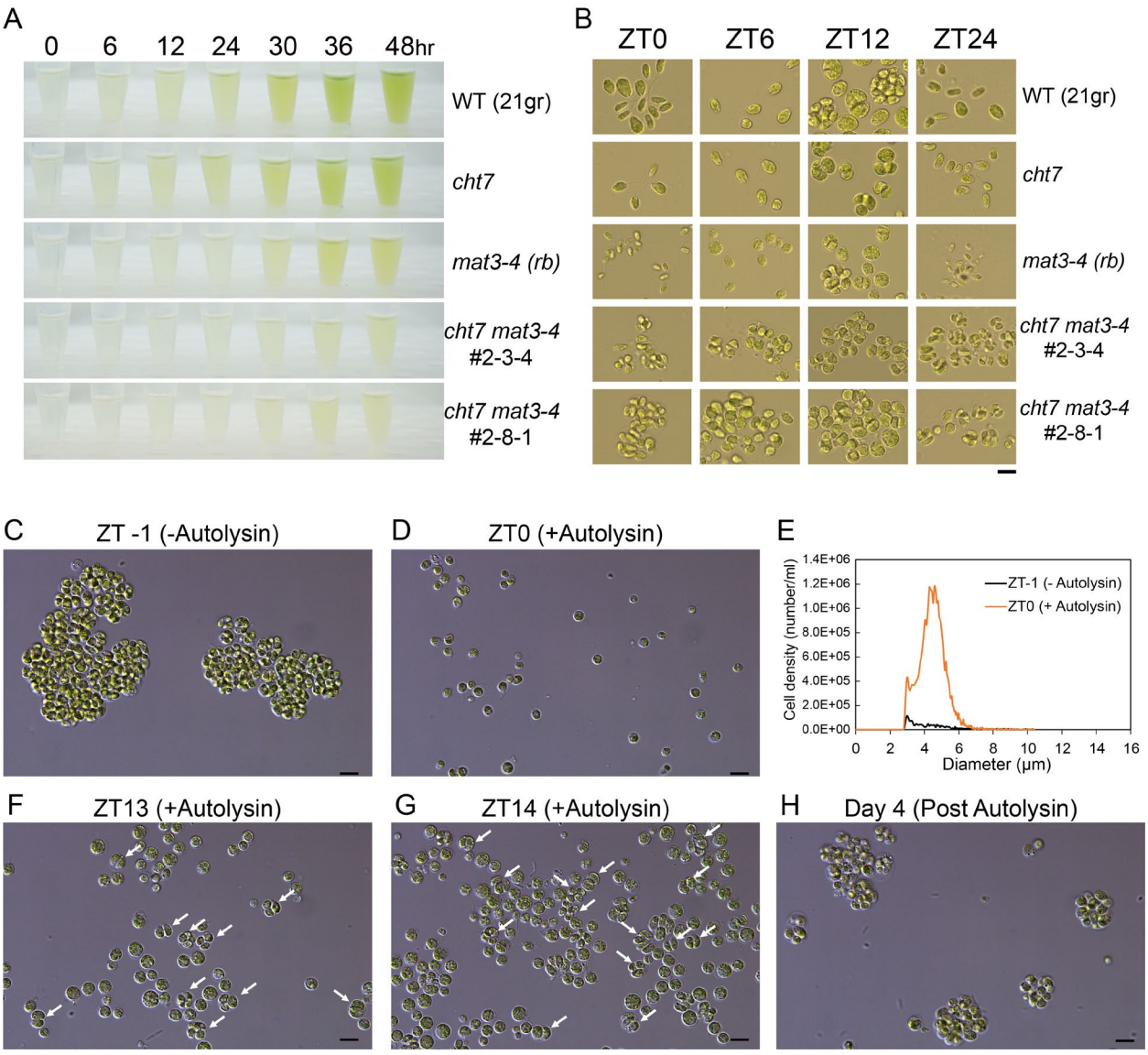


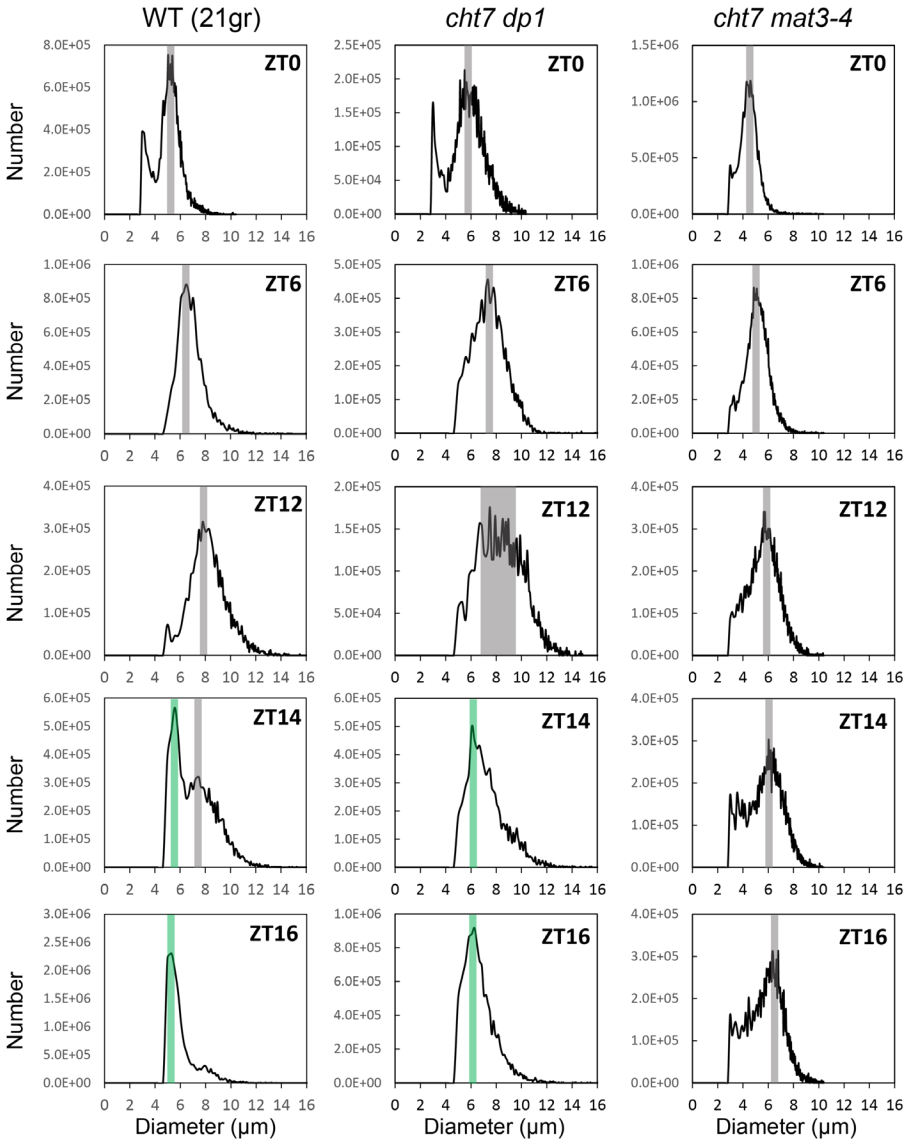
Figure 4.4. Phenotypic analyses of *cht7 mat3-4* during N-replete synchronous growth.

(A) Growth of non-ePBR cultures of *cht7 mat3-4*. Aliquots from cells grown in flasks were removed for visualization in microfuge tubes. ***(B)*** Microscopic images of ePBR-cultures of *cht7 mat3-4* and PLs. ***(C-H)*** Microscopic images of cells before (ZT-1) and after autolysin treatment (at ZT0) of the ePBR cell culture of *cht7 mat3-4*. ***(E)*** Abundance of single cells averaging 5 microns in diameter prior to autolysin treatment (ZT-1) and after treatment (ZT0). White arrows

Figure 4.4 (cont'd)

indicate dividing cells. ZT(n) samples represent cells under normal N+ conditions, where n = hours after the onset of light. Scalebar = 10 μm.

A



B

Strain	% mitotic at ZT0	maximum % mitotic	ZT## of mitotic peak
21gr	0.3%	18 %	ZT13
cht7 dp1	12.1%	31 %	ZT12
cht7 mat3-4	3.7%	33.1 %	>ZT16

C

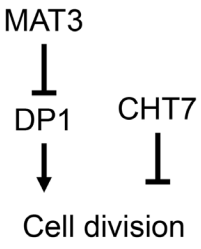


Figure 4.5. Cell cycle analysis of *cht7 dp1* and *cht7 mat3-4* after autolysin treatment during N-replete synchronous growth.

Figure 4.5 (cont'd)

(A) Cell-size distribution of 21gr, *cht7 dp1*, and *cht7 mat3-4* cells at different stages of the cell cycle during N-replete synchronous growth with ePBR. Cell peaks are highlighted in grey (mother cells) or green (daughter cells). X-axis: diameter (μm). Y-axis: cell number. ***(B)*** Microscopic analysis of mitotic activity performed in James Umen's lab. ***(C)*** A hypothesis of the regulatory roles of CHT7 and RB in cell division. Modified from Li et al., 2016. For the ZT(n) samples, n = hours after the onset of light.

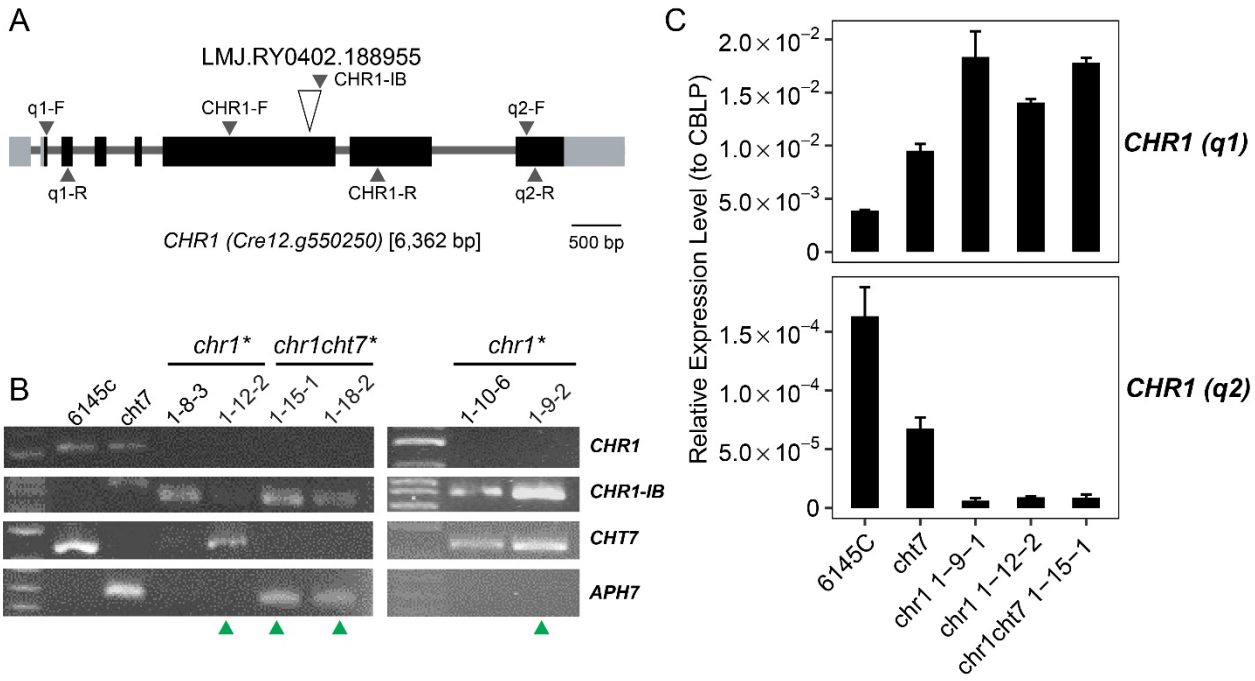


Figure 4.6. Construction and characterization of *chr1* and *chr1 cht7*.

(A) Gene structure of *CHR1 (Cre12.g550250)*. Exons (Black) and UTRs (Grey) are labeled. The insertion site of *chr1* (LMJ.RY0402.188955) from the CLiP mutant library is indicated as a white triangle. Binding sites of primers used in ***(B-C)*** are labeled as black triangles. ***(B)*** PCR validation of the candidate strains (asterisks) of *chr1, chr1 cht7,* and parental lines (PLs). Gene markers are listed below. *chr1*: CHR1 and CHR1-IB (insertion) *cht7*: *CHT7* and *APH7* (*Aminoglycoside Phosphotransferase Gene 7*, insertion marker for paromomycin resistance). Confirmed genotypes are labeled with green triangles. ***(C)*** RT-qPCR analysis of the gene encoding CHR1 in *chr1, chr1 cht7,* and parental lines (PLs) grown under normal conditions.

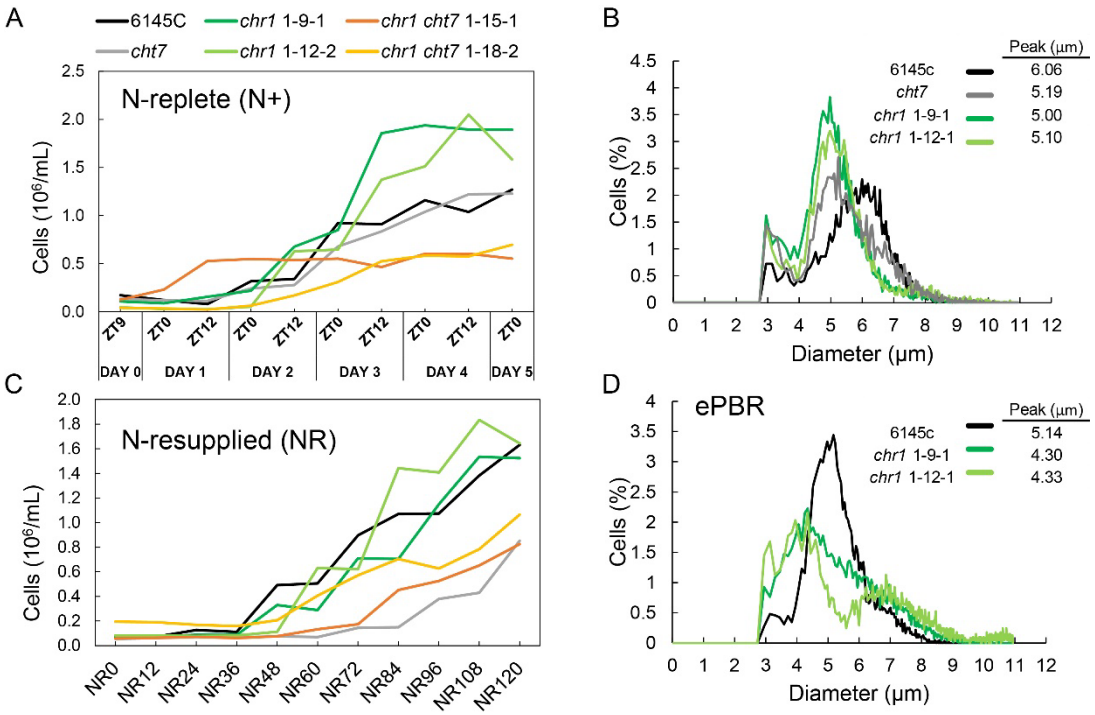


Figure 4.7. Cell growth of *chr1* and *chr1 cht7* under N-replete and N-resupplied conditions. ***(A)*** Growth curves of non-ePBR cell cultures of *chr1*, *chr1 cht7*, and PLs (6145c and *cht7*) during N-replete photoautotrophic growth. ZT(n) = n hours after the onset of light. ***(B)*** Cell size measurement of ZT0 cells of 6145c, *cht7*, and *chr1* in ***(A)***. ***(C)*** Growth curves of N-resupplied *chr1*, *chr1 cht7*, and PLs following N deprivation for 48 hours under photoautotrophic conditions. NR(n) = n hours of N resupply. ***(D)*** Cell size measurement of 6145c, *cht7*, and *chr1* cells collected at ZT0 during N-replete synchronous growth with ePBR. Strains are color labeled as shown above the graph in A.

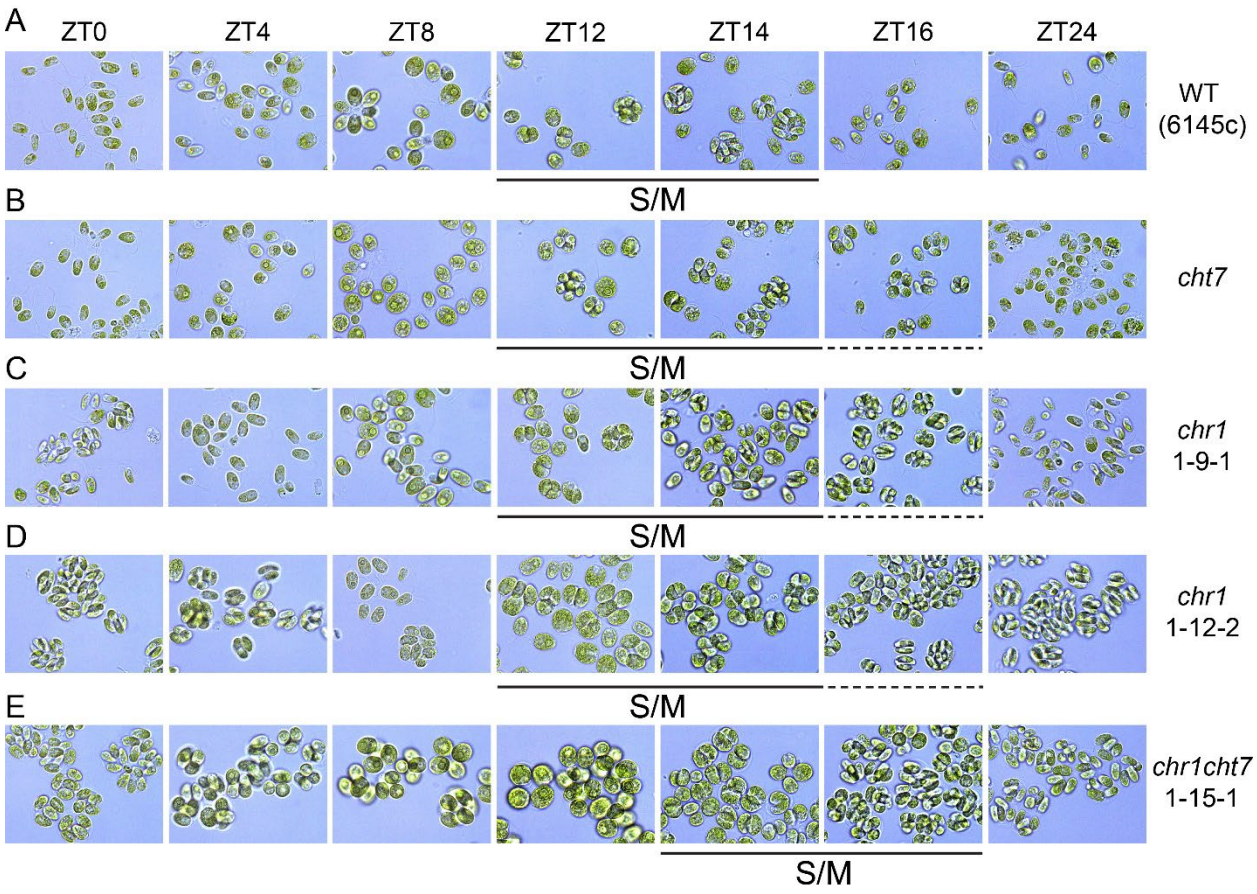


Figure 4.8. Cell cycle progression of *chr1* and *chr1 cht7* of synchronous growth.

Microscopic images of 6145c, *cht7*, *chr1*, *and chr1 cht7* cells at different stages of the cell cycle during N-replete synchronous growth with ePBR. The sampling times with observed S/M-phase activities are indicated. ZT(n) = n hours after the onset of light.

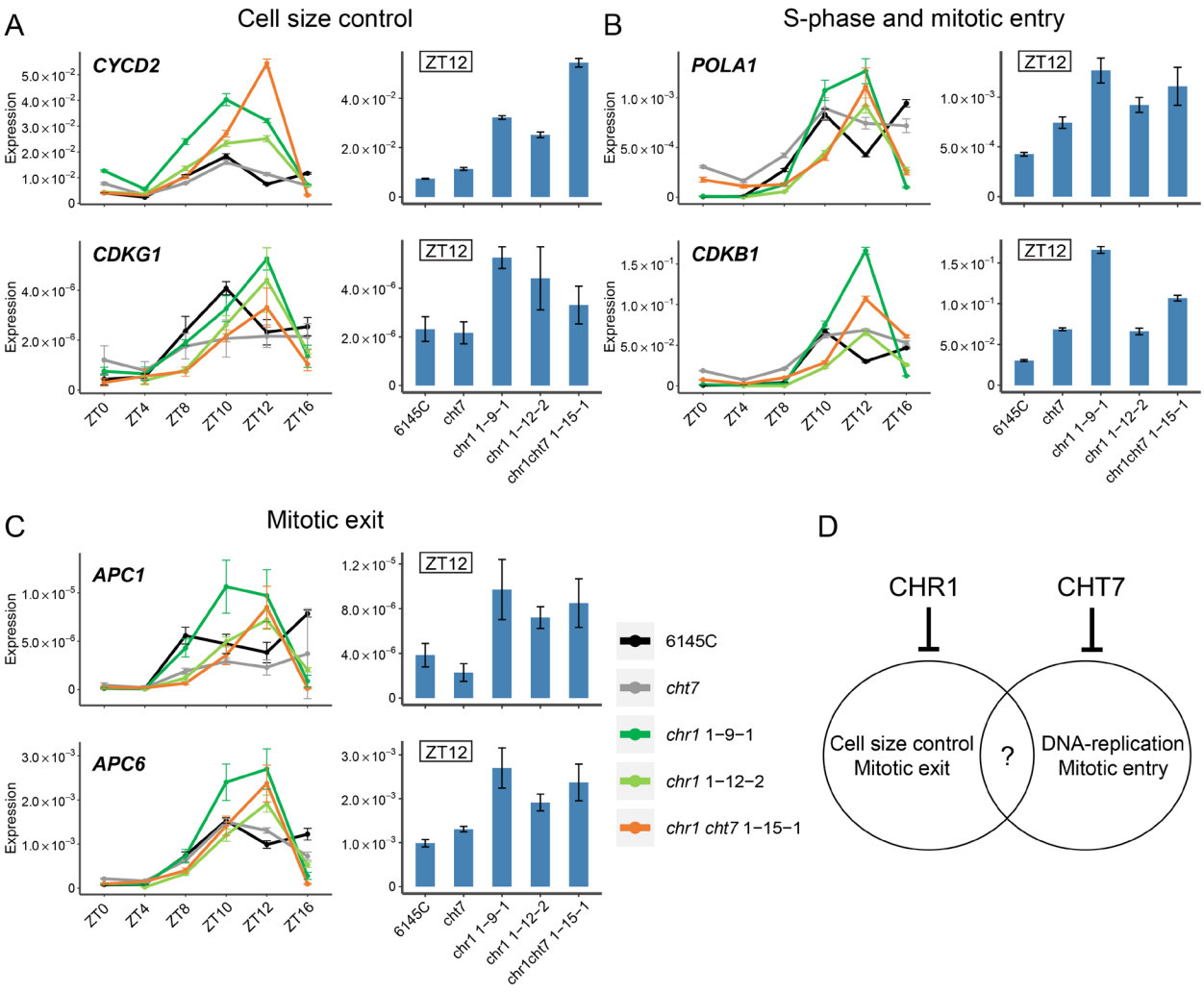


Figure 4.9. Gene expressions of cell cycle markers in *chr1* and *chr1 cht7* cells of synchronous growth.

RT-qPCR analyses of representative cell cycle genes for ***(A)*** Cell size control, ***(B)*** S-phase and mitotic entry, and ***(C)*** Mitotic exit in ePBR cell cultures of 6145c (black), *cht7* (grey), *chr1* (#1-9-1 and 1-12-2, green and light green), *and chr1 cht7* (#1-15-1, orange) during N-replete synchronous growth. ***(D)*** Hypothesized gene regulatory networks of CHR1 and CHT7. ZT(n) = n hours after the onset of light. Left panel: ZT0 to ZT16. Right panel: ZT12 only. *POLA2, DNA Polymerase α2*; *CYCD2, cyclin D2*; *CDKB1/G1: Cyclin-Dependent Kinase B1/G1*; *APC 2/6, Anaphase Promoting Complex 2/6*.

REFERENCES

REFERENCES

Atkins, K.C., and Cross, F.R. (2018). Interregulation of CDKA/CDK1 and the plant-specific cyclin-dependent kinase CDKB in control of the Chlamydomonas cell cycle. Plant Cell 30:429-446.

de Carpentier, F., Lemaire, S.D., and Danon, A. (2019). When unity is strength: The strategies used by Chlamydomonas to survive environmental stresses. Cells 8.

Fang, S.C., de los Reyes, C., and Umen, J.G. (2006). Cell size checkpoint control by the retinoblastoma tumor suppressor pathway. PLoS Genet 2:e167.

Jaenicke, L., Kuhne, W., Spessert, R., Wahle, U., and Waffenschmidt, S. (1987). Cell-wall lytic enzymes (autolysins) of Chlamydomonas reinhardtii are (hydroxy)proline-specific proteases. Eur J Biochem 170:485-491.

Li, Y., Liu, D., Lopez-Paz, C., Olson, B.J., and Umen, J.G. (2016). A new class of cyclin dependent kinase in Chlamydomonas is required for coupling cell size to cell division. Elife 5:e10767.

Lin, Y.T., Takeuchi, T., Youk, B., Umen, J., Sears, B.B., and Benning, C. (2022). Chlamydomonas CHT7 is involved in repressing DNA replication and mitotic genes during synchronous growth. G3 Genes|Genomes|Genetics, 12.

Marceau, A.H., Felthousen, J.G., Goetsch, P.D., Iness, A.N., Lee, H.W., Tripathi, S.M., Strome, S., Litovchick, L., and Rubin, S.M. (2016). Structural basis for LIN54 recognition of CHR elements in cell cycle-regulated promoters. Nat Commun 7:12301.

Matsuo, T., Kuramoto, H., Kumazaki, T., Mitsui, Y., and Takahashi, T. (2012). LIN54 harboring a mutation in CHC domain is localized to the cytoplasm and inhibits cell cycle progression. Cell Cycle 11:3227-3236.

Olson, B.J., Oberholzer, M., Li, Y., Zones, J.M., Kohli, H.S., Bisova, K., Fang, S.C., Meisenhelder, J., Hunter, T., and Umen, J.G. (2010). Regulation of the Chlamydomonas cell cycle by a stable, chromatin-associated retinoblastoma tumor suppressor complex. Plant Cell 22:3331-3347.

Sadasivam, S., and DeCaprio, J.A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. Nat Rev Cancer 13:585-595.

Sadasivam, S., Duan, S., and DeCaprio, J.A. (2012). The MuvB complex sequentially recruits B-Myb and FoxM1 to promote mitotic gene expression. Genes Dev 26:474-489.

Sueoka, N. (1960). Mitotic replication of deoxyribonucleic acid in *Chlamydomonas reinhardi*. Proc Natl Acad Sci U S A 46:83-91.

Takeuchi, T., Lin, Y.T., Fekaris, N., Umen, J., Sears, B.B., and Benning, C. (2020a). Modulation of CHT7 complexes during light/dark- and nitrogen-mediated life cycle transitions of Chlamydomonas. Plant Physiol 184:1762-1774.

Takeuchi, T., Sears, B.B., Lindeboom, C., Lin, Y.T., Fekaris, N., Zienkiewicz, K., Zienkiewicz, A., Poliner, E., and Benning, C. (2020b). Chlamydomonas CHT7 is required for an effective quiescent state by regulating nutrient-responsive cell cycle gene expression. Plant Cell 32:1240-1269.

Tsai, C.H., Warakanont, J., Takeuchi, T., Sears, B.B., Moellering, E.R., and Benning, C. (2014). The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in Chlamydomonas. Proc Natl Acad Sci U S A 111:15833-15838.

Tulin, F., and Cross, F.R. (2015). Cyclin-dependent kinase regulation of diurnal transcription in Chlamydomonas. Plant Cell 27:2727-2742.

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

CHAPTER 5
Conclusions and outlook

Oil produced by microalgae has long been considered an advantageous biofuel feedstock to replace traditional fossil fuels. In comparison with conventional oil crops like soybean and oil palm, microalgae have a much higher annual oil yield per land area (Fig. 1.2) (Barry et al., 2016), which is due to the rapid life cycle and higher oil content per cell in some algae. However, the limited biomass accumulation during the induction of lipid biosynthesis has become an impediment to the industrial-scale production of algal biofuels (Pienkos and Darzins, 2009). To address this problem, we would need to understand the fundamental mechanisms connecting cell proliferation and metabolic status.

The decision-making during cell cycle progression (e.g., to grow or to rest; to divide or not) in microalgae follows the general rule of microorganisms and is mainly dependent on the availability of surrounding nutrients (Fig. 1.1). When nutrients are sufficient, the cell undergoes cell division cycles, and when nutrients become limited, it enters a long-term resting state called quiescence (Gray et al., 2004; Tsai et al., 2014). During quiescence, an algal cell typically stops growing and dividing, and diverts its energy flow to storage carbon compounds like starch and the neutral lipid TAG which can be converted into biodiesel (Li et al., 2012; Miller et al., 2010). Hence, manipulating the mechanism underlying the state transition could potentially remove the constraint on cell growth during TAG accumulation. The green alga *Chlamydomonas reinhardtii* provides a unique platform to study metabolic status-dependent cell cycle regulation in microalgae for the following reasons: 1) it accumulates lipids and starch during quiescence, 2) its fundamental metabolic pathways and cell-cycle mechanisms are well-established, providing a solid research foundation 3) its genome is fully sequenced, and the related multi-omics databases are available (Merchant et al., 2007).

The mutant of Compromised Hydrolysis of Triacylglycerols 7 (CHT7) was isolated from a random insertional mutant screen aimed at identifying at strains with altered TAG levels during the transition out of quiescence that was induced by N deprivation (Tsai et al., 2014). Subsequent work on the *cht7* mutant and the CHT7 protein revealed that instead of directly participating in the biosynthesis of TAG, CHT7 is involved in the transcriptional regulation of the cell cycle. This conclusion is based on several key findings. First, CHT7 belongs to the CHC protein family whose roles in controlling S/M-specific genes are well-established in other species (Tsai et al., 2014). Second, CHT7 proteins are located in the nucleus and the transcriptome analysis of *cht7* showed that some cell cycle genes are mis-regulated during quiescence (Takeuchi et al., 2020b; Tsai et al., 2018; Tsai et al., 2014). Third, some N-deprived *cht7* cells exhibit abnormal phenotypes such as multiple nuclei and arrested cytokinesis-like cell morphology (Takeuchi et al., 2020b). Together, these efforts shed light on a CHC protein-mediated cell cycle regulatory pathway in Chlamydomonas and laid the foundation of this thesis work.

Comprehensive analysis of CHT7-mediated cell cycle regulation

In Chapter 2, we conducted a comprehensive yet in-depth analysis to study the impacts of CHT7 on different stages of the cell cycle. We grew the *cht7* mutant and two control strains, wild type (WT) and the complementation line *CHT7-HA::cht7*, under autotrophic synchronous conditions in photobioreactors with turbidostatic control. Cells cultures above were grown under three growth conditions determined by the availability of nitrogen (N), and each condition represents a unique state of the cell cycle: the normal cell division cycle under the N-replete (N+) condition, the entry into quiescence under the N-deprived (ND) condition, and the exit from quiescence under N-resupplied (NR) condition. In addition, samples of each condition were collected at different time

points representing different stages of the cell cycle. For the N+ condition, samples were harvested at ZT0 (G0 stage), ZT6 (G1 stage), ZT10 and ZT12 (S/M phase), where ZTn represents n hours after the onset of light (ZT: Zeitgeber time). Timepoints of ND and NR conditions are indicated in Fig.2.2A (Lin et al., 2022).

DNA staining coupled with flow cytometry revealed that the fraction of cells containing a ≥ 2C DNA content at non-cell dividing timepoints was larger in *cht7* than in WT and *CHT7-HA::cht7* (Fig. 2.1A-B). We also observed that compared to controls, some *cht7* cells during the G1 phase were round and non-flagellated, and some were even dividing (Fig. 2.1C-D). The abnormal cell morphology and DNA content of the *cht7* mutant suggested an early entry into the S/M phase during the normal cell division cycle. Next, we conducted an RNA-sequencing analysis to globally characterize the effects of CHT7 on gene expression at different stages of the cell cycle. We found that during the non-S/M stages, there were around 1,000 differentially expressed genes (DEGs) constantly induced in the *cht7* mutant compared to WT and *CHT7-HA::cht7* (Fig. 2.3). Most of these genes are expressed during the S/M phase in WT and *CHT7-HA::cht7*. For example, some genes involved in DNA replication and mitosis were derepressed in *cht7* during G0, G1, and quiescence (Fig. 2.4 and 2.5). The mis-regulated mitotic transcriptome in *cht7* resembles the regulatory target of Cyclin-Dependent Kinases CDKA1 and CDKB1 (Tulin and Cross, 2015), which raised the possibility that CHT7 might act in the same pathway of mitotic CDKs. Furthermore, we found that the expression of genes encoding CYCB1, CDKB1, and CYCA1 is also up-regulated in *cht7* during non-cell dividing stages, implying that CHT7 might transcriptionally activate CDKB1, which in turn induces the mitotic transcriptome (Fig. 2.6A). On the other hand, genes encoding CDKA1, the subunits of the Anaphase Promoting Complex (APC), and the cell-size controlling CDKG1 and Retinoblastoma (RB) complex are not affected in the

cht7 mutant (Fig. 2.6C-E; also refer to Chapter 1 for details). Accordingly, we conclude that CHT7 is involved in the negative regulation of several S/M-specific processes, which might counteract mitotic CDKA1-stimulated cell division.

Identification of CHT7-associating regulatory motifs

Data from previous and current transcriptome analyses (Lin et al., 2022; Takeuchi et al., 2020b; Tsai et al., 2014) suggests that CHT7 is a general gene repressor, which prompted us to elucidate the molecular mechanisms of CHT7-mediated gene regulation. However, previous attempts at finding CHT7-bound DNA motifs and CHT7-associated transcription factors (TFs) were unsuccessful (Takeuchi et al., 2020b). In Chapter 3, I developed an alternative approach to search for potential regulatory motifs in promoters of the *cht7*-DEGs identified from the transcriptome data, especially those significantly derepressed during the non-cell dividing time points (Cluster 4, Fig. 3.1 and 3.2). The promoter region in this study is defined as the transcription start site (TSS) +/- 250 bp. To detect DNA motifs specifically enriched in *cht7*-DEGs, a control gene group whose expression was not significantly altered in *cht7* was included in the motif analysis (Fig. 3.1). Consequently, parallel searches with two independent programs identified the enrichment of a motif 5'-AGCTGTCR-3', namely Motif 1, in promoters of C4-*cht7*-DEGs (Fig. 3.3). Interestingly, Motif 1 exhibited a strong local preference of 0-100 bp downstream of the TSS in *cht7*-DEGs compared to the control group, suggesting that this motif is likely a location-specific regulatory element (Fig. 3.3D). A second motif search on Motif 1-containing C4-*cht7*-DEGs further revealed the presence of Motif 2 (5'-GGATCCA-3') located 101-150 bp 5' of Motif 1 (Fig. 3.4). Further analyses of the two motifs indicated that Motif 1 is likely a cis-regulatory element (CRE) recognized by CHT7-associated TFs. Furthermore, the 5' presence of Motif 2 relative to Motif 1

might further enhance the gene repression by the CHT7 protein complex (Fig. 3.5 and 3.6). Notably, the sequence of Motif 1 resembles established CREs recognized by several TALE-class homeodomain (HD)-containing proteins (Fig. 3.7) (Berger et al., 2008; Jolma et al., 2013; Laughon, 1991). Results from protein analyses of HDs of Chlamydomonas and other species indicated a Chlamydomonas homeoprotein named HDG1 is most likely to be the Motif 1-binding TF (Fig. 3.7) (Lee et al., 2008). To test this hypothesis, future studies will include the characterization of the *hdg1* mutant and probing of the CHT7-HDG1 interaction. A HDG1-based ChIP-qPCR experiment probing for Motif 1-containing promoters will also verify this assumption.

Exploring the relationship of *cht7* to other cell cycle regulators

Previous studies have established a role for the CHT7 complex in repressing S/M transcriptional programs during G1 and G0 phases (Lin et al., 2022; Takeuchi et al., 2020b; Tsai et al., 2018). A similar function of CHC proteins in other organisms has been observed (Matsuo et al., 2012; Schmit et al., 2009). In many cases, CHC proteins are associated with RB or RB-like protein complexes to repress genes for cell division during the non-S/M phase (Kobayashi et al., 2015; Sadasivam and DeCaprio, 2013). However, our previous work showed that CHT7 is predominantly in a complex different from the MAT3 (RB)-E2F-DP1 complex in Chlamydomonas (Takeuchi et al., 2020a). In addition, it has been suggested that the MAT3-regulated pathway is mainly involved in the cell size-dependent modulation of cell division during the S/M phase, which is distinct from other established RB pathways (Cross, 2020). To clarify whether the CHT7-regulated gene network is a part of the RB pathway, we generated the double mutant lacking CHT7 and MAT3 or DP1, namely *cht7 mat3-4* and *cht7 dp1* (Fig. 4.2A to C).

The double mutants were grown under N-replete synchronous conditions, as mentioned earlier. We showed that the reduced transcription of genes encoding mitotic cyclins CYCA1 and CYCB1 in *dp1* could be rescued by an additive mutation of CHT7 (Fig. 4.3A). In addition, the larger cell size of the *dp1* mutant caused by a decreased S/M activity was largely reduced in the *cht7 dp1* double mutant (Fig. 4.2D and 4.2H). On the other hand, the induction of the gene encoding CDKB1 in *mat3-4* and *cht7* was further enhanced in the *cht7 mat3-4* (Fig. 4.3B) double mutant showing additivity. The growth of *cht7 mat3-4*, however, was severely impaired and cells often formed palmelloids, a multi-cell structure often linked to stress responses (Fig. 4.4 and 4.5) (de Carpentier et al., 2019). The additivity apparent in these data suggests that the RB and CHT7 complexes act in two distinct pathways or complexes that down-regulate cell division genes (Fig. 4.5C). The mutations of both pathways might result in replicative stress and damage cell growth. The isolation and characterization of the components of the CHT7 and RB complexes will help to validate this hypothesis.

Lastly, we explored the potential role of a CHC protein named CHC-Related 1 (CHR1) in cell cycle regulation. Like other S/M-specific genes, the gene encoding CHR1 is mainly expressed at ZT10 and ZT12 and is derepressed in *cht7* (Fig. 4.1). The *chr1* mutant increased cell density faster than WT and exhibited a smaller cell size during N-replete synchronous growth (Fig. 4.7). In addition, genes encoding CYCD2 and CDKG1, regulators of the cell division activity during S/M, were induced in both *chr1* and *chr1 cht7* at ZT10 and ZT12, which was not seen in *cht7*. These findings support the observation of prolonged S/M in *chr1* and delayed entry of S/M in *chr1 cht7* cells under synchronous growth conditions (Fig. 4.8 and 4.9). We conclude that CHR1 is a novel regulator of the cell cycle in Chlamydomonas, which might control a gene network not targeted by the CHT7 complex (Fig. 4.9D). RNA-sequencing analysis of the transcriptional

regulation in *chr1* during different cell cycle stages will help to clarify this point. Generation of the *chr1* complementation line will also help to confirm the role of CHR1 in cell cycle regulation.

Novel insights into cell cycle regulation of Chlamydomonas

The rhythmic regulation of CDKs and cyclins drives the cell cycle progression in eukaryotes. The transcriptional switches of various cell cycle events are the primary phosphorylation targets of the cyclin-bound CDKs to coordinate different cell cycle stages. In animals, a classical switch is the RB-DP-E2F complex which mediates the G1/S transition. The activity of RB is attenuated by CDK2/CYCE and CDK4/CYCD, resulting in the release of transcription activators E2F1/2/3 and the initiation of the S-phase (Giacinti and Giordano, 2006; Sadasivam and DeCaprio, 2013). During the G0 phase, RB-like proteins p130 and p107 form the DREAM complex with DP, E2F4, and the MuvB core complex to repress cell cycle-dependent genes (Sadasivam and DeCaprio, 2013). The DNA-binding MuvB core plays a critical role in guiding other DREAM subunits to the promoter of target genes. During the G2/M transition, the MuvB core dissociates from the DREAM complex and recruits BMYB and FOXM1 transcription factors to activate mitotic gene expression (Sadasivam and DeCaprio, 2013). The DNA binding specificity of the MuvB core is largely contributed by the CXC protein subunit, LIN54 (Matsuo et al., 2012; Schmit et al., 2009). Together, RB, RB-like, and MuvB core in animals play essential roles in mediating gene expression corresponding to different phases of the cell cycle. In plants, RB-related proteins RBRs have been shown to play a similar role in cell cycle regulation by interacting with DP and E2Fs during the G1/S transition (Zluhan-Martinez et al., 2020). Like animals, the activity of the RB-DP-E2F complex in plants is regulated by CYCD-CDKA (Boniotti and Gutierrez, 2001; Nakagami et al., 2002). The MuvB core is not found in plants, while a DREAM-like complex consisting of the CXC protein TCX5, MYB3R3, and the RBR-DP-E2F

complex has been identified in Arabidopsis and mediates the transition between G2 and M phases (Kobayashi et al., 2015).

In Chlamydomonas, the RB homolog, MAT3, also form a stable, DNA-binding complex with DP and E2F (Olson et al., 2010; Umen and Goodenough, 2001). However, the Chlamydomonas MAT3-DP-E2F complex has a minimal function in cell cycle regulation compared to its homologs in animals and seed plants (Cross, 2020). The most substantial evidence is that mutants of MAT3 and DP in Chlamydomonas are viable and able to maintain the cell cycle (Fang et al., 2006). In contrast, the Arabidopsis null mutant of RBR is gametophytic lethal (Ebel et al., 2004). Furthermore, studies have shown that in Chlamydomonas, the transcriptional activation of the S/M phase is mainly controlled by CDKA1 and B1, and the MAT3 complex is involved in cell size control by regulating the cell division activity during the S/M phase (Cross, 2020; Li et al., 2016). Taken together, the reduced importance of the RB complex in cell cycle regulation of Chlamydomonas suggests that the CHT7 complex might be a microalgae-specific regulatory mechanism that evolved as an alternative to RB (G1/S) and DREAM (G0)-mediated pathways.

Understanding cell cycle regulation towards enhancing biofuel feedstock production in microalgae

This thesis presents a comprehensive analysis of the regulation of genes during the Chlamydomonas cell cycle. We used approaches such as photobioreactor-synchronized cell cultures and bioinformatics tools to study when and how the CHT7 complex mediates the progression of the cell cycle through transcriptional repression. A search of regulatory motifs suggests the potential involvement of a homeoprotein in the CHT7-mediated cell cycle regulatory pathway. The relationship between CHT7 and other cell cycle regulators, including a novel CHC

protein, was also explored. We will continue to detail the function of CHC proteins and their associating TFs in Chlamydomonas, especially the mechanistic details of how these complexes regulate genes and affect cell cycle progression.

Knowledge gained from this research could potentially benefit the development of biofuel feedstocks from microalgae. For example, the removal of inhibitory signals of cell division could potentially increase biomass accumulation under normal growth conditions. The same strategy could also be applied to uncouple the negative correlation between TAG accumulation and cell growth during nutrient-limited conditions. However, it is important to note that the artificial manipulation of the cell cycle might interfere with the well-balanced C/N metabolic pathways and damage the cell, as cell division is highly energy-consuming (Salazar-Roa and Malumbres, 2017). As a result, it is critical to integrate knowledge in cellular energy generation and utilization when developing fast-growing strains with high lipid productivity. While many of the specific details remain to be worked out, this thesis points toward new players as a basis for a novel strategy for developing biofuel feedstock from microalgae by engineering of cell cycle regulation in microalgae.

REFERENCES

REFERENCES

Barry, A., Wolfe, A., English, C., Ruddick, C., and Lambert, D. (2016). 2016 National Algal Biofuels Technology Review. 2016 National Algal Biofuels Technology Review.

Berger, M.F., Badis, G., Gehrke, A.R., Talukder, S., Philippakis, A.A., Pena-Castillo, L., Alleyne, T.M., Mnaimneh, S., Botvinnik, O.B., Chan, E.T., et al. (2008). Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. Cell 133:1266-1276.

Boniotti, M.B., and Gutierrez, C. (2001). A cell-cycle-regulated kinase activity phosphorylates plant retinoblastoma protein and contains, in Arabidopsis, a CDKA/cyclin D complex. Plant J 28:341-350.

Cross, F.R. (2020). Regulation of multiple fission and cell-cycle-dependent gene expression by CDKA1 and the RB-E2F pathway in Chlamydomonas. Curr Biol 30:1855-1865 e1854.

de Carpentier, F., Lemaire, S.D., and Danon, A. (2019). When unity is strength: The strategies used by Chlamydomonas to survive environmental stresses. Cells 8.

Ebel, C., Mariconti, L., and Gruissem, W. (2004). Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. Nature 429:776-780.

Fang, S.C., de los Reyes, C., and Umen, J.G. (2006). Cell size checkpoint control by the retinoblastoma tumor suppressor pathway. PLoS Genet 2:e167.

Giacinti, C., and Giordano, A. (2006). RB and cell cycle progression. Oncogene 25:5220-5227.

Gray, J.V., Petsko, G.A., Johnston, G.C., Ringe, D., Singer, R.A., and Werner-Washburne, M. (2004). "Sleeping beauty": quiescence in *Saccharomyces cerevisiae*. Microbiol Mol Biol Rev 68:187-206.

Jolma, A., Yan, J., Whitington, T., Toivonen, J., Nitta, K.R., Rastas, P., Morgunova, E., Enge, M., Taipale, M., Wei, G., et al. (2013). DNA-binding specificities of human transcription factors. Cell 152:327-339.

Kobayashi, K., Suzuki, T., Iwata, E., Magyar, Z., Bogre, L., and Ito, M. (2015). MYB3Rs, plant homologs of Myb oncoproteins, control cell cycle-regulated transcription and form DREAM-like complexes. Transcription 6:106-111.

Laughon, A. (1991). DNA binding specificity of homeodomains. Biochemistry 30:11357-11367.

Lee, J.H., Lin, H., Joo, S., and Goodenough, U. (2008). Early sexual origins of homeoprotein heterodimerization and evolution of the plant KNOX/BELL family. Cell 133:829-840.

Li, X., Moellering, E.R., Liu, B., Johnny, C., Fedewa, M., Sears, B.B., Kuo, M.H., and Benning, C. (2012). A galactoglycerolipid lipase is required for triacylglycerol accumulation and survival following nitrogen deprivation in Chlamydomonas reinhardtii. Plant Cell 24:4670-4686.

Li, Y., Liu, D., Lopez-Paz, C., Olson, B.J., and Umen, J.G. (2016). A new class of cyclin dependent kinase in Chlamydomonas is required for coupling cell size to cell division. Elife 5:e10767.

Lin, Y.T., Takeuchi, T., Youk, B., Umen, J., Sears, B.B., and Benning, C. (2022). Chlamydomonas CHT7 is involved in repressing DNA replication and mitotic genes during synchronous growth. G3 Genes|Genomes|Genetics, 12.

Matsuo, T., Kuramoto, H., Kumazaki, T., Mitsui, Y., and Takahashi, T. (2012). LIN54 harboring a mutation in CHC domain is localized to the cytoplasm and inhibits cell cycle progression. Cell Cycle 11:3227-3236.

Merchant, S.S., and Prochnik, S.E., and Vallon, O., and Harris, E.H., and Karpowicz, S.J., and Witman, G.B., and Terry, A., and Salamov, A., and Fritz-Laylin, L.K., and Marechal-Drouard, L., et al. (2007). The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318:245-250.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., et al. (2010). Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. Plant Physiol 154:1737-1752.

Nakagami, H., Kawamura, K., Sugisaka, K., Sekine, M., and Shinmyo, A. (2002). Phosphorylation of retinoblastoma-related protein by the cyclin D/cyclin-dependent kinase complex is activated at the G1/S-phase transition in tobacco. Plant Cell 14:1847-1857.

Olson, B.J., Oberholzer, M., Li, Y., Zones, J.M., Kohli, H.S., Bisova, K., Fang, S.C., Meisenhelder, J., Hunter, T., and Umen, J.G. (2010). Regulation of the Chlamydomonas cell cycle by a stable, chromatin-associated retinoblastoma tumor suppressor complex. Plant Cell 22:3331-3347.

Pienkos, P.T., and Darzins, A. (2009). The promise and challenges of microalgal-derived biofuels. Biofuel Bioprod Bior 3:431-440.

Sadasivam, S., and DeCaprio, J.A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. Nat Rev Cancer 13:585-595.

Salazar-Roa, M., and Malumbres, M. (2017). Fueling the cell division cycle. Trends in Cell Biology 27:69-81.

Schmit, F., Cremer, S., and Gaubatz, S. (2009). LIN54 is an essential core subunit of the DREAM/LINC complex that binds to the cdc2 promoter in a sequence-specific manner. FEBS J 276:5703-5716.

Takeuchi, T., Lin, Y.T., Fekaris, N., Umen, J., Sears, B.B., and Benning, C. (2020a). Modulation of CHT7 complexes during light/dark- and nitrogen-mediated life cycle transitions of Chlamydomonas. Plant Physiol 184:1762-1774.

Takeuchi, T., Sears, B.B., Lindeboom, C., Lin, Y.T., Fekaris, N., Zienkiewicz, K., Zienkiewicz, A., Poliner, E., and Benning, C. (2020b). Chlamydomonas CHT7 is required for an effective quiescent state by regulating nutrient-responsive cell cycle gene expression. Plant Cell 32:1240-1269.

Tsai, C.H., Uygun, S., Roston, R., Shiu, S.H., and Benning, C. (2018). Recovery from N deprivation is a transcriptionally and functionally distinct state in Chlamydomonas. Plant Physiol 176:2007-2023.

Tsai, C.H., Warakanont, J., Takeuchi, T., Sears, B.B., Moellering, E.R., and Benning, C. (2014). The protein Compromised Hydrolysis of Triacylglycerols 7 (CHT7) acts as a repressor of cellular quiescence in Chlamydomonas. Proc Natl Acad Sci U S A 111:15833-15838.

Tulin, F., and Cross, F.R. (2015). Cyclin-dependent kinase regulation of diurnal transcription in Chlamydomonas. Plant Cell 27:2727-2742.

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

Zluhan-Martinez, E., Perez-Koldenkova, V., Ponce-Castaneda, M.V., Sanchez, M.P., Garcia-Ponce, B., Miguel-Hernandez, S., Alvarez-Buylla, E.R., and Garay-Arroyo, A. (2020). Beyond What Your Retina Can See: Similarities of Retinoblastoma Function between Plants and Animals, from Developmental Processes to Epigenetic Regulation. Int J Mol Sci 21.