Winter 12-4-2017

Small RNA-Dependent Gene Silencing in the Green Alga *Chlamydomonas reinhardtii*: Functions and Mechanisms

EUN JEONG KIM
University of Nebraska-Lincoln, sassyej@gmail.com

SMALL RNA-DEPENDENT GENE SILENCING IN THE GREEN ALGA CHLAMYDOMONAS REINHARDTII: FUNCTIONS AND MECHANISMS

by

Eun Jeong Kim

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Biological Sciences

(Genetics, Cellular and Molecular Biology)

Under the Supervision of Professor Heriberto Cerutti

Lincoln, Nebraska

December, 2017

SMALL RNA-DEPENDENT GENE SILENCING IN THE GREEN ALGA CHLAMYDOMONAS REINHARDTII: FUNCTIONS AND MECHANISMS

Eun Jeong Kim, Ph.D.

University of Nebraska, 2017

Advisor: Heriberto Cerutti

Small RNAs (sRNAs), ~20-30 nucleotides in length, are non-coding RNAs that play essential roles in the regulation of gene expression in eukaryotes. They lead to inactivation of cognate sequences at the post-transcriptional level via a variety of mechanisms involved in translation inhibition and/or RNA degradation.

In the Chlorophyta *Chlamydomonas reinhardtii*, however, the molecular machinery responsible for sRNA-mediated translational repression remains unclear. To address the mechanisms of translation inhibition by sRNA, we have isolated an RNAi defective mutant (*Mut26*), which contains a deletion of the gene encoding the homolog of CCR4 in Chlamydomonas. We investigated the expression of both an exogenous siRNA target and endogenous miRNA target. Additionally, the pattern of poly(A) tailing in diagnostic mRNAs was examined with the G/I tailing assay and CCR4 partner proteins were identified through affinity purification. Our overall results are consistent with the role of CCR4 in sRNA-dependent translational repression without target mRNA degradation in Chlamydomonas.

The biological function(s) of miRNAs in responses to nutrient deprivation in *Chlamydomonas reinhardtii* were also explored. Transcriptome analysis using cells grown under various trophic conditions revealed that several miRNAs were differentially

expressed, but their predicted targets showed no changes in transcript abundance. Collective evidence suggests that miRNAs may not play an essential role in endogenous gene regulation in Chlamydomonas.

FOR IAN

PREFACE

Chapter 2 has been published in *Bioresource Technology* (**Eun-Jeong Kim**, Xinrong Ma, Heriberto Cerutti). Biore.Tech. 184, 23-32 (2015).

Chapter 3 has been published in *Scientific reports* (**Eun-Jeong Kim** *, Adam Voshall *, Xinrong Ma, Tomohito Yamasaki, Etsuko N. Moriyama & Heriberto Cerutti). Sci. Rep. 7, 5462 (2016). *These authors contributed equally to this work.

Chapter 4 has been published in *Methods in Cell Biology* (**Eun-Jeong Kim**, Heriberto Cerutti. Targeted Gene silencing by RNA interference in Chlamydomonas. Meth. in Cell Bio. 93:99-110 (2009).

TABLE OF CONTENTS

List of Figures viii

List of Table x

CHAPTER 1 – INTRODUCTION

1. RNA interference

A. History of RNA interference (RNAi) pathway 1

B. Argonaute protein at the core of RNAi 3

2. Small RNA-mediated translation repression

A. Translation inhibited by sRNAs in Chlamydomonas 5

B. CCR4-NOT and translation repression 7

3. References 9

CHAPTER 2 - GENE SILENCING IN MICROALGAE: MECHANISMS AND BIOLOGICAL ROLES

Abstract 15

Introduction 16

Results 19

Conclusion and perspective 41

References 41

CHAPTER 3 - MIRNA IN THE ALGA CHLAMYDOMONAS REINHARDTII ARE NOT PHYLOGENETICALLY CONSERVED AND

PLAY A LIMITED ROLE IN RESPONSES TO NUTRIENT DEPRIVATION

Abstract 54

Introduction 55

Results 57

Discussion 71

Methods 80

References 87

CHAPTER 4 - TARGETED GENE SILENCING BY RNA INTERFERENCE IN CHLAMYDOMONAS

Abstract 97

Introduction 98

Materials and Methods 102

Results and Discussion 106

Summary 112

References 113

CHAPTER 5 – THE CHARACTERIZATION OF A CCR4-LIKE PROTEIN IN CHLAMYDOMONAS REINHARDTTI

Abstract 119

Introduction 120

Results 122

Discussion 129

Materials and Methods 131

References 134

CHAPTER 6 – CONCLUSION 141

References 143

List of Figures

Chapter 1

Figure 1.1: RNA silencing pathways 2

Figure 1.2: Overall architecture of Argonaute from the three domains of life. 4

Figure 1.3: The functions of mRNAs in land plants, metazoans and the green alga Chlamydomonas. 6

Figure 1.4: miRNA-mediated translational repression in animals. 8

Chapter 2

Figure 2.1: Maximum likelihood tree showing the phylogenetic relationship among KMT1 proteins. 21

Figure 2.2: Maximum likelihood tree showing the phylogenetic relationship among KMT6 proteins. 23

Figure 2.3: Maximum likelihood tree showing the phylogenetic relationship among Dnmt1/MET1 proteins. 27

Figure 2.4: Maximum likelihood tree showing the phylogenetic relationship among Argonaute proteins. 34

Chapter 3

Figure 3.1: Comparison of miRNAs identified in Chlamydomonas cells grown under various nutritional deprivation conditions. 60

Figure 3.2: Northern blot analysis of miRNA expression in Chlamydomonas

cells grown under the denoted trophic conditions. 62

Figure 3.3: Comparison of predicted miRNA targets in Chlamydomonas cells grown under various nutritional deprivation conditions. 64

Figure 3.4: Growth and survival of Chlamydomonas cells subjected to various nutritional deprivation conditions. 66

Figure 3.5: Growth and survival of Chlamydomonas cells subjected to various nutritional deprivation conditions. 69

Figure 3.6: Distribution of miRNAs based on their average expression levels for C. reinhardtii. 74

Figure 3.7: Comparison of the number of predicted cleavage targets for lowly expressed and highly expressed miRNAs. 75

Figure 3.8: Relationship between miRNA expression level and number of predicted targets in Arabidopsis thaliana and Arabidopsis lyrata. 78

Chapter 4

Figure 4.1: TIR–RNAi system for silencing of the EXO gene in Chlamydomonas. 101

Figure 4.2: Silencing of the EXO gene by tandem IR transgenes in transformants selected in the 5-FI-containing medium. 109

Chapter 5

Figure 5.1: The MAA7 protein is increased in the RNAi defective Mut26, containing a deletion of the CCR4 protein. 123

Figure 5.2: Analysis of miRNA targets transcript and protein abundance. 124

Figure 5.3: Analysis of miR912 expression and Ago3 protein abundance. 125

Figure 5.4: Analysis of polyadenylated tail lengths of the MAA7 and ACT1 transcripts of the indicated strains. 127

Figure 5.5: Co-immunoprecipitation of the Flag-tagged-MUT26 associated proteins. 128

List of Tables

Table 2.1: Distribution of core gene silencing components in Archaeplastida microalgae. 20

Table 5.1: Proteins associated with the Flag-tagged CCR4 protein. 129

CHAPTER 1

INTRODUCTION

1. RNA interference

A. History of the RNA interference (RNAi) pathway

In plants, RNA-mediated gene silencing was first observed in 1986, when antisense RNA was found to effectively block the expression of target genes. (Ecker & Davis, 1986). In 1990, co-suppression of gene silencing was demonstrated by two groups, who attempted to overexpress a flower pigment enzyme in the petunia family to obtain darker purple flowers. However, the transgenic plants exhibited wholly or partially white flowers, indicating that endogenous genes and transgenes were downregulated in the white parts of the petunia flower (Krol, Mur, Beld, MOI, & Stuitje, 1990; Napoli, Lemieux, & Jorgensen, 1990). In *C. elegans*, injections of nmy-2 (non-muscle myosin II) antisense RNA into ovaries of adult worms caused embryonic partitioning defects and led to mislocalization of PAR proteins, which are required for the early asymmetrical divisions. This indicated that antisense RNAs could act as gene silencers of the muscle myosin II genes when injected in *C. elegans* (Guo & Kemphues, 1996). This finding led to the discovery that double-stranded RNAs (dsRNA) were the actual initiators of post-transcriptional gene silencing in *C. elegans.* dsRNA-mediated gene silencing was named RNA interference (RNAi) by Fire and Mello, who were later awarded the Nobel Prize in Physiology and Medicine in 2006 (Fire *et al*., 1998). The finding of dsRNAs as a precursor promoted the discovery of an RNase III named Dicer and Argonaute (Ago).

Dicers produce short dsRNA molecules with phosphorylated 5' ends and 2-nucleotide overhangs at the 3' ends. Agos act as silencing effector proteins that target specific mRNAs for repression (Bernstein, Caudy, Hammond, & Hannon, 2001; Bohmert *et al*., 1998) (Figure 1.1).

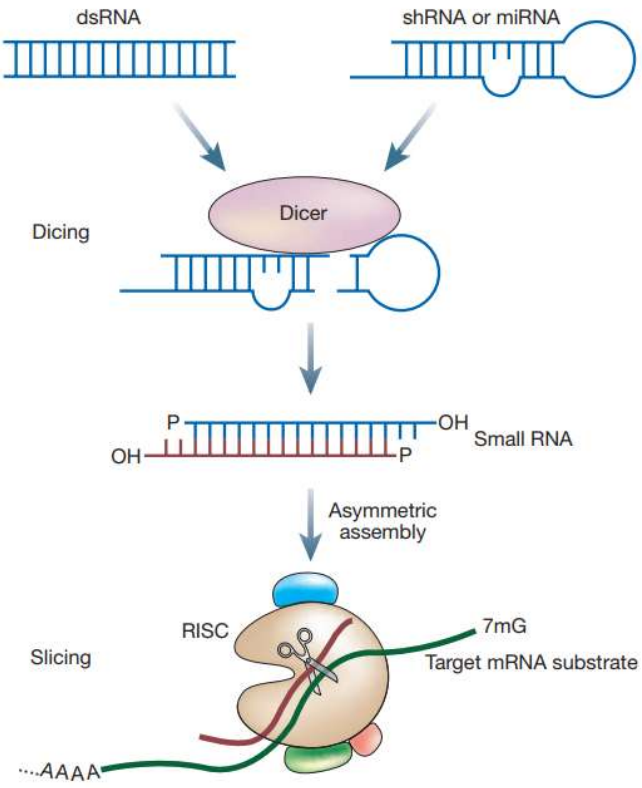

Figure 1.1. [Taken from (Hannon & Rossi, 2004)] RNA silencing pathways. Long dsRNA and miRNA precursors are processed to siRNA/miRNA duplexes by the RNase-III-like enzyme Dicer. These short dsRNAs are subsequently unwound and assembled into effector complexes, RISCs, which can direct RNA cleavage or mediate translational repression. In *D. melanogaster* and *A. thaliana*, specialized Dicer or DLC proteins preferentially process long dsRNA or miRNA precursors. 7mG, 7-methyl guanine; AAAA, poly-adenosine tail; P, 5' phosphate.

Protein structure analyses have elucidated their domain structures and the mechanism of binding with small RNAs (Lau, Potter, Carragher, & MacRae, 2009; Song *et al*., 2003). Both Dicer and Ago are phylogenetically conserved proteins and core

components of the RNAi machinery (David P. Bartel, 2004). To date, the canonical RNAi pathway is the best-characterized mechanism of RNA-mediated gene silencing. Since the discovery of RNAi, gene silencing phenomenon have been investigated in a broad spectrum of eukaryotes, from fungi to humans (Zamore & Haley, 2005), and recently also in prokaryotes, bacteria and archaea (Swarts *et al*., 2014; Zander *et al*., 2017).

B. The argonaute protein at the core of RNAi

Genetic, biochemical, and crystallographic studies have revealed the structure of the core RNAi machinery, especially Argonaute proteins that control the catalytic activity (Carbonell *et al*., 2012; Cenik & Zamore, 2011; Song *et al*., 2003). Ago proteins are highly conserved in eukaryotes, except for yeast (Saccharomyces cerevisiae) that lost their RNAi machinery (Nakanishi, Weinberg, Bartel, & Patel, 2012) (Figure 1.2). Also, some bacteria and archaea possess Ago proteins, but their specific functions remained unknown until 2014 (Tolia & Joshua-Tor, 2007). That year, an Ago of bacterium *Thermus thermophilus* was shown to participate in host defense via DNA interference (Swarts *et al*., 2014). Later, FRET (Fluorescence Resonance Energy Transfer) was used to elucidate the functions of archaean Ago. This technique revealed the cleavage activity and conformational change capabilities of Ago proteins upon the guide strand during the target strand loading in the thermophilic methanogenic archaean, *Methanocaldococcus jannaschii* (Zander *et al*., 2017).

Ago proteins contain a PAZ (PIWI-ARGONAUTE-ZWILLE) domain, which binds the methylated 3'end of sRNAs into a specific pocket and a MID (middle) domain

that interacts with the 5'end of sRNAs (Meister, 2013). The proteins also contain an N domain at the N-terminus and a PIWI (originally named P-element-induced wimpy testis) domain at the C-terminus. Crystallographic analysis of *Tt*Ago (*Thermus thermophiles*) showed that the N domain prevents extension of the guide strand and biochemical studies of *Hs*Ago (*Homo sapiens*) revealed that the N domain facilitates the unwinding of the

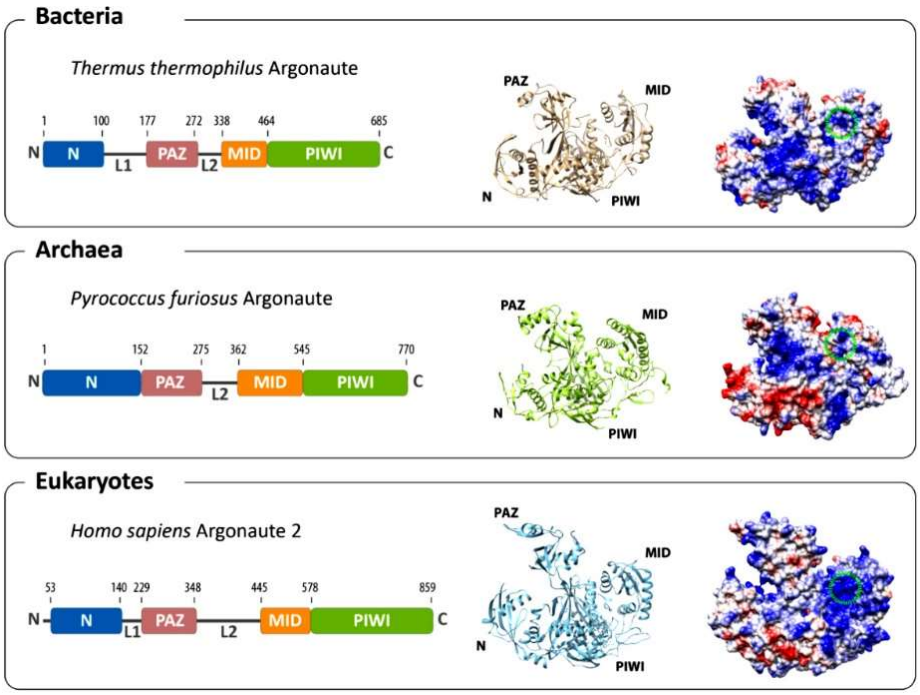

Figure 1.2. [Taken from (Willkomm, Zander, Gust, & Grohmann, 2015)] Overall architecture of Argonaute from the three domains of life. The domain composition (left) and structures (middle) of the bacterial (based on Thermus thermophilus, PDB: 3DLH), the archaeal (based on Pyrococcus furiosus, PDB: 1U04) and the eukaryotic (based on human Argonaute 2, PDB: 4EI3) Argonaute reveal an evolutionarily conserved architecture. Differences can be found in the surface charge distribution of Argonaute proteins (negatively-charged surfaces in red; positively-charged surfaces in blue). The binding pocket for the 5'-end of the guide in the MID domain is highlighted with a green circle.

guide/target duplex during RISC (RNA-Induced Silencing Complex) assembly (Miyoshi, Ito, Murakami, & Uchiumi, 2016). The PIWI domain is structurally similar to RNase H,

which functions as an endonuclease. Additionally, it contains a catalytic triad composed of DDX (where X is D or H) amino acid residues and a recent structural study of *S. pombe* Ago proteins revealed that a fourth residue is essential, turning the catalytic center into a tetrad of DEDX residues (Nakanishi *et al*., 2012; Tolia & Joshua-Tor, 2007). Lastly, Ago structures contain interconnectors L1 and L2 between N and PAZ, and the PAZ and MID domains, respectively (Figure 1.2). They are located on the guide-RNA interfacing surface of the PIWI domain and their conformational changes affect either guide-RNA stability or cleavage activity (Wang *et al*., 2009).

In summary, Agos are evolutionally conserved from all kingdoms of life as well as specialized proteins for RNAi pathways. For instance, there are 10 Agos in *Arabidopsis thaliana*, 5 Agos in *Caenorhabditis elegans*, and 3 Agos in *Chlamydomonas reinhartti* (Hock & Meister, 2008).

2. Small RNA-mediated translational repression

A. Translation inhibited by sRNAs in Chlamydomonas

Small RNAs (sRNAs) regulate gene expression at both the transcriptional and post-transcriptional level in multiple ways, including DNA/histone methylation, chromosome segregation, RNA stability, and protein translation (Carthew & Sontheimer, 2009). Post-transcriptional gene regulation, especially translational repression mechanism(s), has been gaining attention because this repression affects protein production without altering the steady level of transcripts. This sophisticated molecular mechanism(s) is still unexplained in various aspects. For instance, unknown components include how inhibition is regulated at the translation initiation step, deadenylation

followed by then decapping and mRNA degradation, proteolysis (degradation of the nascent peptide), and how is blocked (slowed elongation or ribosome drop-off) (Filipowicz, Bhattacharyya, & Sonenberg, 2008).

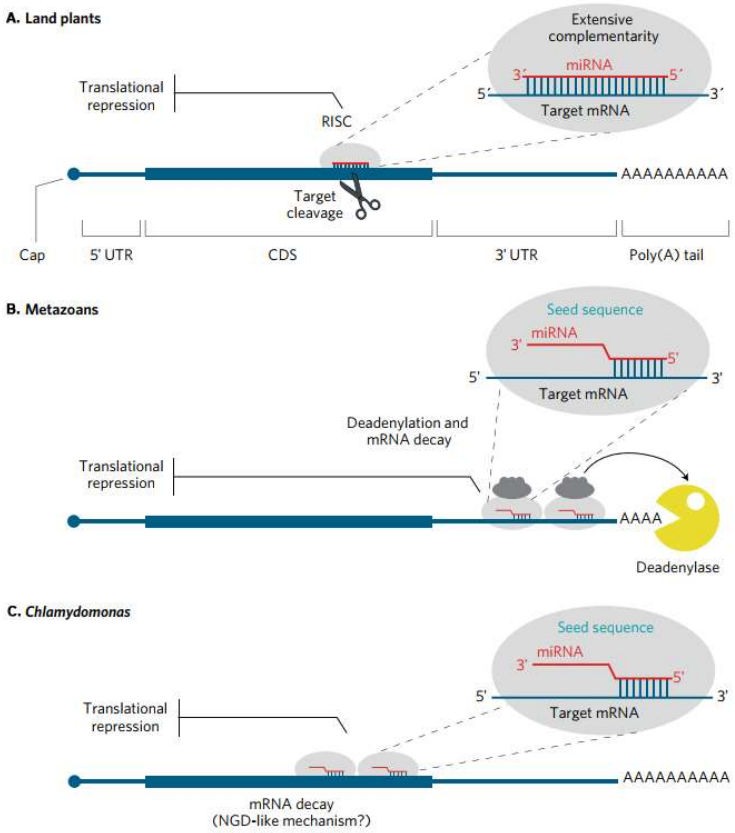

Figure 1.3. [Taken from (Iwakawa & Tomari, 2017)] The functions of microRNAs in land plants, metazoans, and the green alga Chlamydomonas. (A) In land plants, miRNA targets are silenced by endonucleolytic cleavage and translational repression generally via a single miRNA binding site, which is extensively complementary to the cognate miRNA. (B) In metazoans, miRNAs promote translational repression as well as deadenylation and subsequent mRNA decay by guiding RISCs to recruit deadenylases and other decay activators to target mRNAs 3'UTR via limited sequence complementarity through the so-called seed region (C) In Chlamydomonas, miRNAs regulate endogenous target mRNAs by RNA degradation and/or translational repression via binding to the CDs with seed complementary.

Plant miRNAs were initially shown to cleave target mRNAs, whereas animal miRNAs targeted transcripts for translational repression (Iwakawa & Tomari, 2017). However, recent genetic and biochemical studies have revealed that both plants and animals seem to operate in both modes of gene silencing action, including cleavage and translational repression. Even with this commonality though, miRNA-mediated gene silencing may still be mechanistically different in each kingdom (Iwakawa & Tomari, 2015). Moreover, the role of sRNAs in gene silencing mechanisms is unclear in unicellular alga like Chlamydomonas (Ma *et al.*, 2013). Current findings suggest that Chlamydomonas sRNA mechanisms are unique when compared to higher plants and animals (Chung, Deery, Groen, Howard, & Baulcombe, 2017; Iwakawa & Tomari, 2017) (Figure 1.3). Chapter 5 will attempt to shed light on some aspects of translational repression mechanisms in *Chlamydomonas reinhardtii*.

B. CCR4/NOT and translational repression

To date, many studies in different organisms (although mostly animals) have suggested that miRNAs inhibit the initiation step of translation (Flamand *et al*., 2016; Nishimura & Fabian, 2016). Yet, it is still unclear how miRNAs precisely repress translation initiation. Recent biochemical studies discovered that miRISC blocks the assembly of the eIF4F translation initiation factor in Drosophila (Fukaya, Iwakawa, & Tomari, 2014) and miRISC promotes the release of both eIF4AI and eIF4II from the target mRNA (Fukao *et al*., 2014).

In the past few years, three major mechanisms of translation inhibition by miRISC have been proposed, including GW182-mediated PABP displacement,

recruitment of translational repressors through GW182, and dissociation of eIF4A from the cap-binding complex eIF4F (Iwakawa & Tomari, 2015). GW182-mediated translational repression and mRNA decay have generally been reported in animal systems, where miRISC directly recruits the GW182 protein bound to the CCR4/NOT complex, resulting in deadenylation, decapping, and target mRNA degradation (Jonas & Izaurralde, 2015)

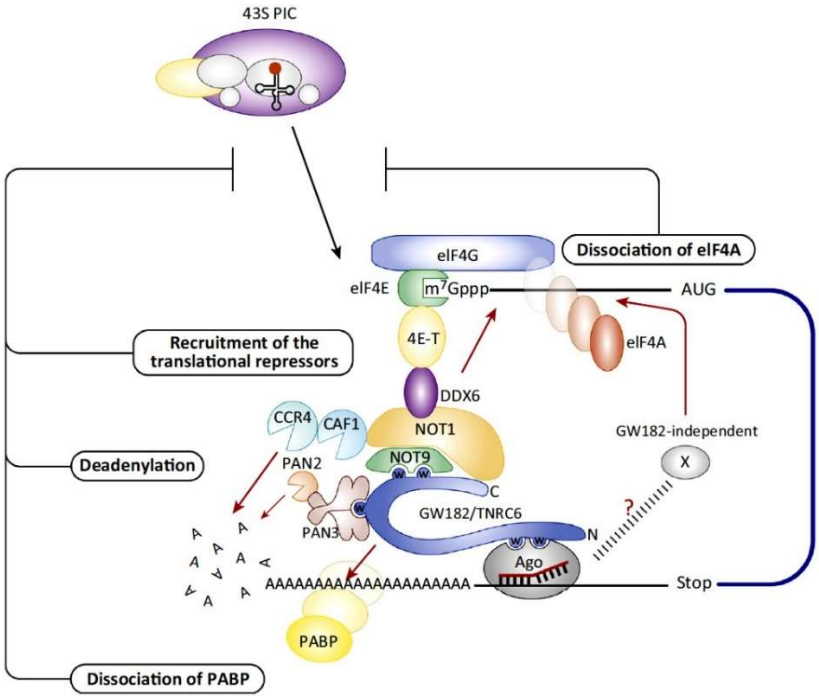

Figure 1.4. [Taken from (Iwakawa & Tomari, 2015)] miRNA-Mediated Translational Repression in Animals. Through the CCR4–NOT complex, GW182 recruits downstream translational repressors (e.g., DDX6 or 4E-T) onto target mRNAs, but it remains that they block unclear which step(s) in translation. miRNAs can also repress translation via displacement of the ATP-dependent RNA helicase eIF4A from the cap-binding complex eIF4F, which blocks the 43S PIC recruitment or ribosomal scanning in vitro. This mode of miRNA action does not require GW182 in Drosophila. Deadenylation as well as the displacement of poly(A)-binding protein (PABP) through GW182 and CCR4–NOT also contribute to the overall miRNA-mediated translational repression.

(Figure 1.4). The CCR4/NOT complex is a multi-subunit protein complex evolutionarily conserved across eukaryotes, which regulates several aspects of gene expression, including transcription initiation, control of messenger RNA elongation, deadenylation and subsequent degradation of mRNA, translation, and even protein decay (Chapat and Corbo, 2014). Strikingly, despite lacking GW182 homolog proteins, plant miRNAs also showed a disproportional decrease in the protein accumulation of their targets, suggesting that miRNA-mediated translational repression also occurs in plant systems (Iwakawa & Tomari, 2015; Voshall *et al*., 2017).

In summary, plant miRNA-mediated translational repression could possibly be involved in an unexplored mechanism, such as a GW182-independent pathway or the participation of a cofactor connected with the translation initiation complexes (Figure 1.4). Indeed, Drosophila Ago1 was shown to repress translation primarily by an ATP-dependent shortening of the poly(A) tail of its mRNA target in a GW182-dependent way. In contrast, Drosophila Ago2 inhibited the cap function in a GW182-independent way (Iwasaki, Kawamata, & Tomari, 2009).

3. References

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell, 116*, 281-297.

Baulcombe, D. (2004). RNA silencing in plants. *Nature, 431*, 356-363.

Bernstein, E., Caudy, A. A., Hammond, S. M., & Hannon, G. J. (2001). Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature, 409*.

Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., & Benning, C. (1998). AGO1 defines a novel locus of Arabidopsis controlling leaf development. *The EMBO J., 17*(1), 170-180.

Carbonell, A., Fahlgren, N., Garcia-Ruiz, H., Gilbert, K. B., Montgomery, T. A., Nguyen, T., Carrington, J. C. (2012). Functional analysis of three Arabidopsis ARGONAUTES using slicer-defective mutants. *Plant Cell, 24*(9), 3613-3629. doi:10.1105/tpc.112.099945

Carthew, R. W., & Sontheimer, E. J. (2009). Origins and Mechanisms of miRNAs and siRNAs. *Cell, 136*(4), 642-655. doi:10.1016/j.cell.2009.01.035

Cenik, E. S., & Zamore, P. D. (2011). Argonaute proteins. *Curr Biol, 21*(12), R446-449. doi:10.1016/j.cub.2011.05.020.

Cerutti, H. (2003). RNA interference: traveling in the cell and gaining functions? *Trends in Genetics, 19*(1), 39-46. doi:10.1016/s0168-9525(02)00010-0.

Chung, B. Y., Deery, M. J., Groen, A. J., Howard, J., & Baulcombe, D. C. (2017). Endogenous miRNA in the green alga Chlamydomonas regulates gene expression through CDS-targeting. *Nat Plants*. doi:10.1038/s41477-017-0024-6.

Hannon, Gregory J., Rossi, John J. (2004) Unlocking the potential of the human genome with RNA interference. *Nature*, 431, 371-378.

Ecker, J. R., & Davis, R. W. (1986). Inhibition of gene expression in plant cells by expression of antisense RNA. *Proc. Natl. Acad. Sci. USA, 83*(5372-5376).

Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet, 9*(2), 102-114. doi:10.1038/nrg2290

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., & Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature, 391*.

Flamand, M. N., Wu, E., Vashisht, A., Jannot, G., Keiper, B. D., Simard, M. J., . . . Duchaine, T. F. (2016). Poly(A)-binding proteins are required for microRNA-mediated silencing and to promote target deadenylation in C. elegans. *Nucleic Acids Res, 44*(12), 5924-5935. doi:10.1093/nar/gkw276

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., . . . Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. *Mol Cell, 56*(1), 79-89. doi:10.1016/j.molcel.2014.09.005

Fukaya, T., Iwakawa, H. O., & Tomari, Y. (2014). MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. *Mol Cell, 56*(1), 67-78. doi:10.1016/j.molcel.2014.09.004

Guo, S., & Kemphues, K. J. (1996). A non-muscle myosin required for embryonic polarity in Caenorhabditis elegans. *Nature, 382*, 455-458.

Hammond, S. M., Bernstein, E., Beach, D., & Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. *Nature, 404*, 293-296.

Hock, J., & Meister, G. (2008). The Argonaute protein family. *Genome Biol, 9*(2), 210. doi:10.1186/gb-2008-9-2-210

Iwakawa, H. O., & Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol, 25*(11), 651-665. doi:10.1016/j.tcb.2015.07.011

Iwakawa, H. O., & Tomari, Y. (2017). Silencing messages in a unique way. *Nat Plants*. doi:10.1038/s41477-017-0028-2

Iwasaki, S., Kawamata, T., & Tomari, Y. (2009). Drosophila argonaute1 and argonaute2 employ distinct mechanisms for translational repression. *Mol Cell, 34*(1), 58-67. doi:10.1016/j.molcel.2009.02.010

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet, 16*(7), 421-433. doi:10.1038/nrg3965

Krol, A. R. v. d., Mur, L. A., Beld, M., MOI, J. N. M., & Stuitje, A. R. (1990). Flavonoid Genes in Petunia: Addition of a Limited Number of Gene Copies May Lead to a Suppression of Gene Expression. *The Plant Cell, 2*(291-299).

Lau, P. W., Potter, C. S., Carragher, B., & MacRae, I. J. (2009). Structure of the human Dicer-TRBP complex by electron microscopy. *Structure, 17*(10), 1326-1332. doi:10.1016/j.str.2009.08.013

Ma, X., Kim, E. J., Kook, I., F. Voshall, A. Moriyama, & E. Cerutti, H. (2013). Small Interfering RNA-Mediated Translation Repression Alters Ribosome Sensitivity to Inhibition by Cycloheximide in Chlamydomonas reinhardtii. *Plant Cell, 25*(3), 985-998. doi:10.1105/tpc.113.109256

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet, 14*(7), 447-459. doi:10.1038/nrg3462

Meister, G., Landthaler, M., Patkaniowska, A., Dorsett, Y., Teng, G., & Tuschl, T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol Cell, 15*(2), 185-197. doi:10.1016/j.molcel.2004.07.007

Miyoshi, T., Ito, K., Murakami, R., & Uchiumi, T. (2016). Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat Commun, 7*, 11846. doi:10.1038/ncomms11846

Nakanishi, K., Weinberg, D. E., Bartel, D. P., & Patel, D. J. (2012). Structure of yeast Argonaute with guide RNA. *Nature, 486*(7403), 368-374. doi:10.1038/nature11211

Napoli, C., Lemieux, C., & Jorgensen, R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes Ín trans. *The Plant Cell*.

Nishimura, T., & Fabian, M. R. (2016). Scanning for a unified model for translational repression by microRNAs. *EMBO J, 35*(11), 1158-1159. doi:10.15252/embj.201694324

Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., . . . Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol, 10*(12), 1026-1032. doi:10.1038/nsb1016

Swarts, D. C., Makarova, K., Wang, Y., Nakanishi, K., Ketting, R. F., Koonin, E. V., . . . van der Oost, J. (2014). The evolutionary journey of Argonaute proteins. *Nat Struct Mol Biol, 21*(9), 743-753. doi:10.1038/nsmb.2879

Tolia, N. H., & Joshua-Tor, L. (2007). Slicer and the argonautes. *Nat Chem Biol, 3*(1), 36-43. doi:10.1038/nchembio848

Voshall, A., Kim, E. J., Ma, X., Yamasaki, T., Moriyama, E. N., & Cerutti, H. (2017). miRNAs in the alga Chlamydomonas reinhardtii are not phylogenetically

conserved and play a limited role in responses to nutrient deprivation. *Sci Rep, 7*(1), 5462. doi:10.1038/s41598-017-05561-0

Wang, Y., Juranek, S., Li, H., Sheng, G., Wardle, G. S., Tuschl, T., & Patel, D. J. (2009). Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes. *Nature, 461*(7265), 754-761. doi:10.1038/nature08434

Willkomm, S., Zander, A., Gust, A., & Grohmann, D. (2015). A prokaryotic twist on argonaute function. *Life (Basel), 5*(1), 538-553. doi:10.3390/life5010538

Wilson, R. C., & Doudna, J. A. (2013). Molecular mechanisms of RNA interference. *Annu Rev Biophys, 42*, 217-239. doi:10.1146/annurev-biophys-083012-130404

Zamore, P. D., & Haley, B. (2005). Ribo-gnome: the big world of small RNAs. *Science, 309*(5740), 1519-1524. doi:10.1126/science.1111444

Zamore, P. D., Tuschl, T., Sharp, P. A., & Bartel, D. P. (2000). RNAi. *Cell, 101*(1), 25-33. doi:10.1016/s0092-8674(00)80620-0

Zander, A., Willkomm, S., Ofer, S., van Wolferen, M., Egert, L., Buchmeier, S., . . . Grohmann, D. (2017). Guide-independent DNA cleavage by archaeal Argonaute from Methanocaldococcus jannaschii. *Nat Microbiol, 2*, 17034. doi:10.1038/nmicrobiol.2017.34

CHAPTER 2

GENE SILENCING IN MICROALGAE: MECHANISMS AND BIOLOGICAL ROLES

Abstract

Microalgae exhibit enormous diversity and can potentially contribute to the production of biofuels and high-value compounds. However, for most species, our knowledge of their physiology, metabolism, and gene regulation is fairly limited. In eukaryotes, gene silencing mechanisms play important roles in both the reversible repression of genes that are required only in certain contexts and the suppression of genome invaders such as transposons. The recent sequencing of several algal genomes is providing insights into the complexity of these mechanisms in microalgae. Collectively, glaucophyte, red, and green microalgae contain the machineries involved in repressive histone H3 lysine methylation, DNA cytosine methylation, and RNA interference. However, individual species often only have subsets of these gene silencing mechanisms. Moreover, current evidence suggests that algal silencing systems function in transposon and transgene repression but their role(s) in gene regulation or other cellular processes remains virtually unexplored, hindering rational genetic engineering efforts.

Introduction

Algae are a diverse group of eukaryotic organisms with important roles in marine, freshwater, and terrestrial ecosystems (Worden and Allen, 2010; Tirichine and Bowler, 2011). The great potential of algae as feedstocks for renewable biofuel and biomaterial production is also gaining recognition (Hu *et al*., 2008; Radakovits *et al*., 2010; Gimpel *et al*., 2013; Leite *et al*., 2013). Microalgae are microscopic organisms capable of harnessing sunlight and CO2 to synthesize useful chemical compounds, such as lipids and carbohydrates, which can be converted into fuels and other bioproducts. However, production of algae-based fuels is technically, but not yet economically, feasible (Lee, 2011; Chisti, 2013). The major economic bottlenecks cited in the literature include microalgae biological productivity, culture systems, crop protection, and harvesting/extraction processes (Hu *et al*., 2008; Chisti, 2013; Gimpel *et al*., 2013; Leite *et al*., 2013).

For large-scale fuel production reliant on algal photosynthesis key objectives will be achieving high productivity per unit of area, environmental (biotic and abiotic) stress tolerance, ease of harvesting and extraction, and a biomass profile optimized for biofuel conversion (Griffiths and Harrison, 2009; Radakovits *et al*., 2010). However, identifying in nature microalgal strains simultaneously endowed with all these traits has proven difficult (Hu *et al*., 2008; Griffiths and Harrison, 2009). Additionally, there has been limited success in increasing biomass productivity or oil content in algae by the genetic engineering of individual genes (Radakovits *et al*., 2010; La Russa *et al*., 2012; Gimpel *et al*., 2013), and this limitation emphasizes the importance of comprehending on a genome

scale the metabolic and regulatory networks involved in these processes. Indeed, a significant barrier to advancement is that our knowledge of gene function and regulation is still fairly incomplete in most microalgae (Radakovits *et al*., 2010; Worden and Allen, 2010; Tirichine and Bowler, 2011). In this context, the study of algal gene silencing mechanisms may provide insights into the control of gene expression as well as facilitate the development of tools for rational genetic engineering.

The regulation of gene expression in eukaryotes involves complex mechanisms, operating at the transcriptional and posttranscriptional levels. Chromatin organization modulates the access of regulatory proteins to DNA and influences multiple aspects of transcription and other DNA-related processes (Bannister and Kouzarides, 2011; Ohsawa *et al*., 2013). Eukaryotic genomes are commonly organized into several types of chromatin, with euchromatin consisting of transcriptionally permissive or active domains and heterochromatin being characterized by densely packed silent regions (Casas-Mollano *et al*., 2007; Krauss, 2008; Bannister and Kouzarides, 2011). These functionally and structurally different chromatin states are marked by distinct covalent modifications on the DNA and on specific amino acid residues of the nucleosomal histones (Casas-Mollano *et al*., 2007; Bannister and Kouzarides, 2011; Saze and Kakutani, 2011). For instance, di- or trimethylation of histone H3 lysine 9 (H3K9) or of histone H3 lysine 27 (H3K27) is often associated with silenced chromatin (Krauss, 2008; Shaver *et al*., 2010; Bannister and Kouzarides, 2011; Saze and Kakutani, 2011; Derkacheva and Hennig, 2014). DNA cytosine methylation also plays a role in repression and in some organisms there appears to be a complex interplay between histone tail modifications and DNA

methylation in establishing a silent chromatin structure (Law and Jacobsen, 2010; Saze and Kakutani, 2011; Du *et al*., 2012; Zhong *et al*., 2014).

RNA-directed mechanisms have also been co-opted by evolution to generate a broad spectrum of gene regulatory pathways. RNA-mediated silencing is a conserved process in eukaryotes by which small RNAs (~20–30 nucleotides in length) induce the inactivation of cognate sequences through a variety of mechanisms, including translation inhibition, RNA degradation, and/or transcriptional repression (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Meister, 2013). The function of long double stranded RNAs, as precursors of small RNAs, in triggering gene silencing was initially characterized in Caenorhabditis elegans and termed RNA interference (RNAi) (Fire *et al*., 1998). Yet, in slightly over a decade, RNAi has evolved into a fascinating biological phenomenon intersecting with multiple cellular pathways. Indeed, histone post-translational modifications, DNA cytosine methylation, and RNA-mediated mechanisms impinge on many cellular processes including, besides regulation of gene expression, DNA repair and recombination, chromatin structure, chromosome condensation/stability, as well as the suppression of viruses and transposable elements (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Cerutti *et al*., 2011; Ohsawa *et al*., 2013; Oliver *et al*., 2014). Moreover, gene silencing mechanisms seem to be important for the integration of environmental and intrinsic stimuli in the control of gene expression and their disruption leads to physiological and developmental abnormalities (Bannister and Kouzarides, 2011; Ohsawa *et al*., 2013).

In most algal species, both chromatin-associated and RNA-mediated silencing pathways remain largely uncharacterized, even at the level of identifying crucial gene

factors in the sequenced genomes. This review will examine the existence of key histone lysine methyltransferases, DNA cytosine methyltransferases, and core components of the RNA-mediated silencing machinery in microalgae. However, algae are very diverse phylogenetically (Worden and Allen, 2010; Tirichine and Bowler, 2011) and, to simplify the identification of commonalities, the analysis will be restricted to microalgae in the Archaeplastida eukaryotic supergroup, which includes glaucophytes (Glaucophyta), red algae (Rhodophyta), green algae (Chlorophyta), as well as land plants (Streptophyta) (Table 1). We will also discuss briefly the known or inferred biological role(s) of gene silencing mechanisms in these aquatic organisms. It is anticipated that advances in the basic understanding of gene regulatory mechanisms in microalgae will enable optimization of metabolic pathways of interest through hypothesis-driven genetic engineering strategies.

Results

2.1 H3K9 and H3K27 methyltransferases in microalgae

2.1.1 Phylogenetic analysis and domain organization of histone methyltransferases

The methylation of lysine residues in histones, with the exception of H3K79 methylation, is carried out by enzymes that contain an evolutionary conserved SET domain, named after three Drosophila genes (Su(var)3-9, Enhancer of zeste, and Trithorax) (Casas-Mollano *et al*., 2007; Bannister and Kouzarides, 2011; Huang *et al*., 2011; Derkacheva and Hennig, 2014). The SET domain constitutes the catalytic site of these lysine methyltransferases (KMTs), but flanking sequences, more distant protein

domains, and possibly some cofactors are also important for enzyme activity and specificity (Huang *et al*., 2011; Krishnan *et al*., 2011; Derkacheva and Hennig, 2014). To begin characterizing the occurrence and the role(s) of H3K9 and/or H3K27 methyltransferases in microalgae, we surveyed 14 complete or near-complete algal

Table 2.1. Distribution of core gene silencing components in Archaeplastida microalgae.

Species	Genome size (Mb)	AGO-Piwi	Dicer	RDR	DNA methyltransferases				Histone H3 methyltransferases		References[b]
					Dnmt1	CMT	Others[a]	Dnmt3	KMT1	KMT6	
Glaucophyta											
Cyanophora paradoxa	70.0	0[c]	0	1	0	0	1	0	0	0	9
Rhodophyta											
Galdieria sulphuraria	13.7	0	0	0	0	0	0	1	0	1	11
Cyanidioschyzon merolae	16.5	0	0	0	0	0	0	1	0	1	5
Porphyridium purpureum	19.7	3	1	3	0	0	0	0	0	0	1
Chlorophyta											
Ostreococcus tauri	12.5	0	0	0	0	0	0	0	0	1	4
Ostreococcus lucimarinus	13.2	0	0	0	0	0	0	0	0	1	8
Bathycoccus prasinos	15.1	0	0	0	0	0	0	0	0	1	7
Micromonas sp. RCC299	20.9	1	0	0	0	0	0	0	0	1	13
Micromonas pusilla CCMP1545	21.9	0	0	0	0	0	0	0	1	1	13
Chlorella variabilis NC64A	46.0	1	1	0	1	1	1	0	2	1	2
Chlorella sorokiniana	56.8	1	1	0	1	1	1	0	2	1	Draft
Coccomyxa subellipsoidea	48.8	2	1?[d]	1	0	0	2	0	2	1	3
Chlamydomonas reinhardtii	120.0	3	3	0	3[e]	0	1	0	1	1	6
Volvox carteri	138.0	2	1	0	1	0	0	0	1	1	10
Streptophyta											
Arabidopsis thaliana	125.0	10	4	6	4	3	0	2	10	3	12

a DNA methyltransferases that cannot be clearly categorized (see text for details).
b References: 1, Bhattacharya *et al*., 2013; 2, Blanc *et al*., 2010; 3, Blanc *et al*., 2012; 4, Derelle *et al*., 2006; 5, Matsuzaki *et al*., 2004; 6, Merchant *et al*., 2007; 7, Moreau *et al*., 2012; 8, Palenik *et al*., 2007; 9, Price *et al*., 2012; 10, Prochnik *et al*., 2010; 11, Schönknecht *et al*., 2013; 12, Arabidopsis Genome Initiative, 2000; 13, Worden *et al*., 2009.
c Total number of genes in the genome encoding a certain factor.
d?, predicted protein fairly divergent from canonical model.
e Includes the chloroplast targeted DMT1 methyltransferase and its putative paralog (see text for details).

genomes in the Archaeplastida supergroup for the presence of SETdomain polypeptides (Table 1).

Proteins with conserved SET domains were identified by either BLAST or PSI-BLAST searches of protein and/or translated genomic DNA databases, using as queries

known *Arabidopsis thaliana* or *Homo sapiens* polypeptides containing SET motifs. Since several of the examined genomes are in draft stage, an important caveat in our analyses is that some proteins may be missing from the databases whereas others may have errors in the predicted gene structure. However, with few exceptions, we considered as potential homologs only proteins that exhibited enough sequence similarity to be aligned and used

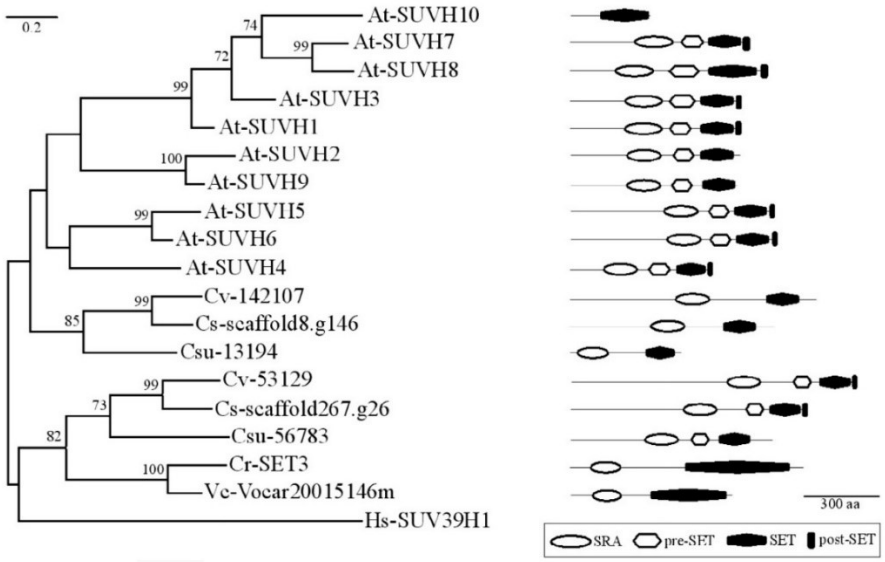

Figure 2.1. Maximum likelihood tree showing the phylogenetic relationship among KMT1 proteins. Sequences corresponding to KMT1 proteins from different organisms were aligned using Muscle and the tree was drawn using the MEGA v6.06 program. Numbers indicate bootstrap values, >60%, based on 1000 pseudoreplicates. Species are designated by a two- or three-letter abbreviation preceding the name of each protein. At, *Arabidopsis thaliana*; Cr, *Chlamydomonas reinhardtii*; Cs, *Chlorella sorokiniana*; Csu, *Coccomyxa subellipsoidea* C-169; Cv, *Chlorella variabilis* NC64A; Hs, *Homo sapiens*; Vc, *Volvox carteri*. Accession numbers of proteins used to draw the tree are: At-SUVH1 to At-SUVH9, AAK28966 to AAK28974, respectively; At-SUVH10, NP_178647; Cr-SET3, XP_001701764; Cs-scaffold8.g146 and Cs-scaffold267.g26, from *Chlorella sorokiniana* draft genomic sequence assembly at the University of Nebraska-Lincoln; Csu-13194, XP_005649968; Csu-56783, XP_005646116; Cv-142107, XP_005850072; Cv-53129, XP_005846419; Hs-SUV39H1, NP_001269095; Vc-Vocar20015146m, XP_002952737. The SMART 7 database was used to identify conserved protein motifs and the domain organization of the proteins is indicated on the right (see text for details).

for phylogenetic tree construction. Interestingly, phylogenetic analysis of the extracted SET-domain proteins revealed that they could be grouped into several distinct classes (see also Huang *et al*., 2011) but we only examined in detail KMT1 and KMT6 homologs (Figures. 1 and 2), following the nomenclature proposed for yeast and metazoan lysine methyltransferases (Allis *et al*., 2007).

Members of the algal KMT1 class, like animal and plant KMT1 proteins, are likely responsible for H3K9 methylation, an epigenetic mark involved in gene silencing and heterochromatin formation (Casas-Mollano *et al*., 2007; Krauss, 2008; Bannister and Kouzarides, 2011; Huang *et al*., 2011). In the examined microalgae, KMT1 homologs appear to be limited to species of the Chlorophyta clade, including organisms in the Trebouxiophyceae (*Chlorella sorokiniana*, *Chlorella variabilis* NC64A, and *Coccomyxa subellipsoidea*) and Chlorophyceae (*Chlamydomonas reinhardtii* and *Volvox carteri*) classes (Table 2.1 and Figure 2.1). *Micromonas pusilla* CCMP1545 also seems to code for a KMT1 related protein (Table 2.1).

Yet, the corresponding gene is located in an island of the genome with no detectable homology to the closely related *Micromonas sp*. RCC299 (data not shown) and the functional significance of the encoded protein is currently unclear. Most algal KMT1 proteins show high sequence similarity to land plant KMT1 polypeptides, both within the SET domain and in the surrounding regions known as the Pre-SET and Post-SET motifs (Figure 2.1). Additionally, all algal sequences contain an SRA (SET and RING associated) domain (Figure 2.1), which recognizes the methylation status of CG and CHH DNA sequences (where H = A, T, or C) (Rajakumara *et al*., 2011). Land plant KMT1 proteins have been reported to fall into several distinct subgroups, indicative of

functional diversification (Casas-Mollano *et al*., 2007; Huang *et al*., 2011). However, our phylogenetic analysis suggests that gene duplication and potential pathway diversification occurred after the divergence of land plants from the lineage leading to green algae (Figure 2.1).

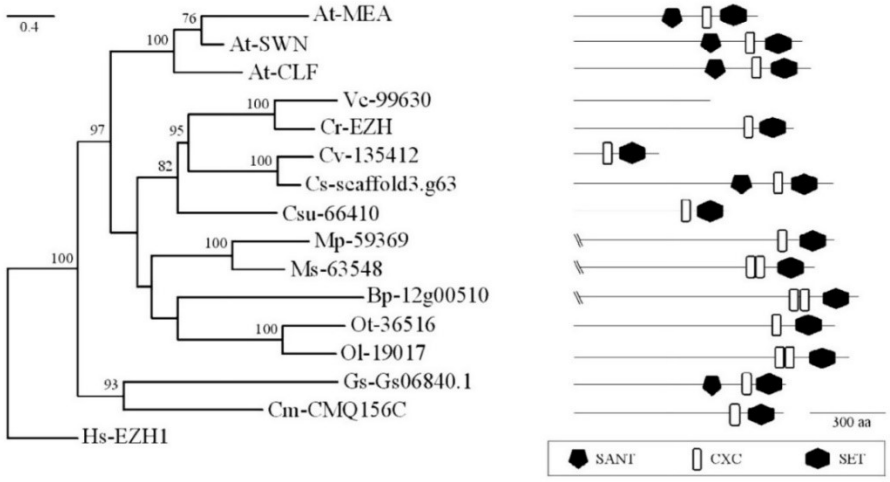

Figure 2.2. Maximum likelihood tree showing the phylogenetic relationship among KMT6 proteins. Sequences corresponding to KMT6 proteins from different organisms were aligned using Muscle and the tree was drawn using the MEGA v6.06 program. Numbers indicate bootstrap values, >60%, based on 1000 pseudoreplicates. Species are designated by a two-letter or three-letter abbreviation preceding the name of each protein, as described in the legend to Figure 2.1, except for: Bp, *Bathycoccus prasinos*; Cm, *Cyanidioschyzon merolae*; Gs, *Galdieria sulphuraria*; Mp, *Micromonas pusilla* CCMP1545; Ms, *Micromonas sp*. RCC299; Ol, *Ostreococcus lucimarinus*; Ot, *Ostreococcus tauri*. Accession numbers of proteins used to draw the tree are: At-CLF, NP_179919; At-MEA, NP_563658; At-SWN, NP_567221; Bp-12g00510, XP_007509828; Cm-CMQ156C, XP_005538096; Cr-EZH, Cre17.g746247 at http://phytozome.jgi.doe.gov; Cs-scaffold3.g63, from Chlorella sorokiniana draft genomic sequence assembly at the University of Nebraska-Lincoln; Csu-66410, PID 66410 at http://genome.jgi.doe.gov/pages/search-for-genes.jsf?organism=Coc_C169_1; Cv-135412, PID 135412 at http://genome.jgipsf.org/ChlNC64A_1/ChlNC64A_1.home.html; Gs-Gs06840.1, XP_005708796; Hs-EZH1, NP_001982; Mp-59369, XP_003059747; Ms-63548, XP_002505475; Ol-19017, XP_001422533; Ot-36516, XP_003082958; Vc-99630, XP_002958369. The SMART 7 database was used to identify conserved protein motifs and the domain organization of the proteins is indicated on the right (see text for details).

A phylogenetic tree of KMT6 histone methyltransferases supports a similar conclusion (Figure 2.2). Members of the algal KMT6 class, like animal and plant KMT6 proteins, are likely responsible for H3K27 methylation and putative components of the evolutionarily conserved Polycomb Repressive Complex 2 (Shaver *et al*., 2010; Bannister and Kouzarides, 2011; Huang *et al*., 2011; Derkacheva and Hennig, 2014). Within the examined microalgae, KMT6 homologs appear to be widely distributed (Table 2.1 and Figure 2.2). Only the Glaucophyta clade seems to lack this type of histone methyltransferases, with the caveat that the draft genome of Cyanophora paradoxa may be incomplete and this observation may have to be revisited upon genome completion and/or the sequencing of additional species in the clade. Most algal KMT6 polypeptides show high sequence similarity to land plant KMT6 proteins, particularly within the SET domain (Figure 2.2). Additionally, almost all algal sequences contain a specific CXC motif (Figure 2.2), an ~ 65-residue cysteine-rich region that has been implicated in DNA binding and shows structural homology to Pre-SET domains (Zheng *et al*., 2012). However, with two exceptions, algal KMT6 proteins lack the SANT (SWI3, ADA2, N-CoR, and TFIIIB DNA-binding) motif that is found in most land plant homologs (Huang *et al*., 2011). Overall our findings are consistent, as previously suggested (Shaver *et al*., 2010), with widespread distribution of KMT6 methyltransferases within algal lineages.

2.1.2 Biological role(s) of H3K9 and H3K27 methylation

By analogy to their function in higher eukaryotes (Casas-Mollano *et al*., 2007; Krauss, 2008; Shaver *et al*., 2010; Bannister and Kouzarides, 2011; Huang *et al*., 2011; Derkacheva and Hennig, 2014), these post-translational histone modifications are likely

involved in transcriptional repression in microalgae. Yet, experimental evidence is almost entirely lacking. To our knowledge, direct determination of H3K9 and H3K27 methylation has only been carried out in *C. reinhardtii*. This alga contains monoand trimethylated H3K9 as well as mono- and dimethylated H3K27 (Casas-Mollano *et al*., 2007; Shaver *et al*., 2010). Additionally, high levels of H3K9 monomethylation have been found associated with transcriptionally repressed transgenes (Casas-Mollano *et al*., 2007; Strenkert *et al*., 2013). Interestingly, RNAi-mediated suppression of Chlamydomonas SET3, a KMT1 methyltransferase (Figure 2.1), reactivated the expression of silenced repetitive transgenic arrays (Casas-Mollano *et al*., 2007).

RNAi-mediated suppression of Chlamydomonas EZH, a KMT6 methyltransferase (Figure 2.2), also led to defects in the silencing of transgenes and retrotransposons as well as to a global increase in histone post-translational modifications associated with transcriptional activity, such as trimethylation of histone H3 lysine 4 and acetylation of histone H4 (Shaver *et al*., 2010). Intriguingly, a C. variabilis NC64A virus, Paramecium bursaria chlorella virus 1 (PBCV-1), has been shown to encode a KMT6 protein, termed vSET, that might be linked to the rapid inhibition of host transcription after viral infection (Wei and Zhou, 2010). In a heterologous system, vSET causes H3K27 methylation and transcriptional repression (Mujtaba *et al*., 2008; Wei and Zhou, 2010). These observations, taken together, are consistent with a role of H3K9 and H3K27 methylation in gene silencing in microalgae, possibly as a means to suppress intragenomic parasites such as transposable elements, as well as in host-virus interactions. However, H3K27 methyltransferases appear to have a much broader taxonomic distribution than H3K9 methyltransferases (Table 1), perhaps implicating

proteins of the KMT6 class in a more essential, evolutionarily conserved regulatory role(s). Nevertheless, assessing the potential involvement of posttranslational histone modifications in specific gene or pathway regulation in microalgae, as demonstrated in land plant and animals (Bannister and Kouzarides, 2011; Derkacheva and Hennig, 2014), will require further studies.

2.2 DNA cytosine methyltransferases in microalgae

2.2.1 Phylogenetic analysis and domain organization of DNA methyltransferases

The DNA of a vast array of organisms contains the modified base 5-methylcytosine, but genomes are methylated in different ways and with different consequences in diverse species (Feng *et al*., 2010; Zemach *et al*., 2010; Law and Jacobsen, 2010; Huff and Zilberman, 2014). Introduction of a methyl group at the C5 position of cytosine is catalyzed by a large family of DNA methyltransferases, including six subfamilies characterized by catalytic domains associated with N-terminal or C-terminal extensions containing distinct motifs (Goll and Bestor, 2005; Ponger and Li, 2005; Huff and Zilberman, 2014). In land plants, DNA cytosine methylation can occur in three sequence contexts namely CG, CHG, and CHH (where H = A, T, or C) and three DNA methyltransferase subfamilies have been implicated in the establishment and/or maintenance of methylation at these sequences (Goll and Bestor, 2005; Law and Jacobsen, 2010). CG methylation is maintained by the Dnmt1/MET1 subfamily of DNA methyltransferases, CHG methylation is mediated by the plant-specific chromomethylases (CMTs), and CHH methylation is introduced by the Dnmt3/DRM (Domains Rearranged Methyltransferase) enzymes (Goll and Bestor, 2005; Law and

Jacobsen, 2010; Zhong *et al*., 2014). Interestingly, in a clear interplay between silencing mechanisms, chromomethylases appear to be targeted to chromosomal regions with H3K9 methylated nucleosomes via their bromo adjacent homology (BAH) and chromo domains (Du *et al*., 2012), whereas DRMs seem to be guided to target loci by small RNAs and certain components of the RNAi machinery (Law and Jacobsen, 2010; Zhong *et al*., 2014).

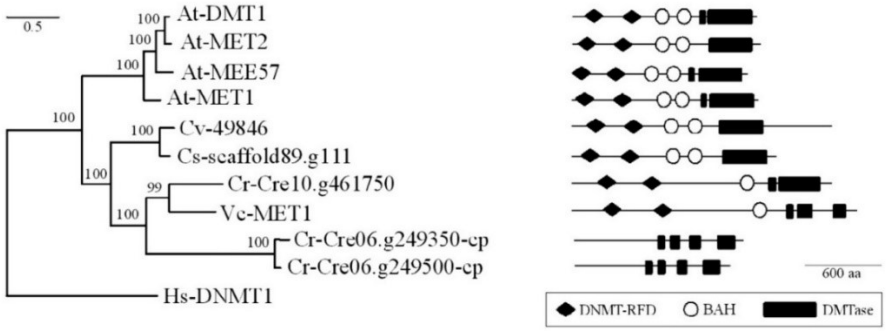

Figure 2.3. Maximum likelihood tree showing the phylogenetic relationship among Dnmt1/MET1 proteins. Sequences corresponding to the DMTase domain from different organisms were aligned using Muscle and the tree was drawn using the MEGA v6.06 program. Numbers indicate bootstrap values, >60%, based on 1000 pseudoreplicates. Species are designated by a two-letter or three-letter abbreviation preceding the name of each protein, as described in the legend to Figure 2.1. Accession numbers of proteins used to draw the tree are: At-MET1, NP_199727; At-DMT1, NP_192638; At-MET2, NP_001190725; At-MEE57, NP_193097; Cr-Cre10.g461750, Cr-Cre06.g249350-cp, and Cr-Cre06.g249500-cp, Cre10.g461750, Cre06.g249350, and Cre06.g249500 at http://phytozome.jgi.doe.gov; Cs-scaffold89.g111, from *Chlorella sorokiniana* draft genomic sequence assembly at the University of Nebraska-Lincoln; Cv-49846, PID 56675 at http://genome.jgi-psf.org/ChlNC64A_1/ChlNC64A_1.home.html; Hs-DNMT1, NP_001370;Vc-MET1, Vocar20014971m at http://phytozome.jgi.doe.gov. The SMART 7 database was used to identify conserved protein motifs and the domain organization of the proteins is indicated on the right (see text for details).

The presence of Dnmt1/MET1, CMT, and Dnmt3/DRM polypeptides was examined in the 14 microalgae with sequenced genomes belonging to the Archaeplastida supergroup, using as queries known A. thaliana or H. sapiens polypeptides (Table 1).

Homologs of Dnmt3/DRM enzymes, which have been implicated in de novo DNA methylation (Goll and Bestor, 2005; Ponger and Li, 2005; Law and Jacobsen, 2010), were found only in the red algae Galdieria sulphuraria and Cyanidioschyzon merolae (Table 1) and with a structural organization similar to the vertebrate enzymes (data not shown). Whereas chromomethylase-related methyltransferases appear to be restricted to the genus Chlorella (Table 1). Dnmt1/MET1 homologs were identified in species of the Trebouxiophyceae (*C. sorokiniana* and *C. variabilis*) and Chlorophyceae (*C. reinhardtii* and *V. carteri*) classes (Table 2.1 and Figure 2.3). The algal Dnmt1/MET1 proteins show high sequence similarity to the land plant polypeptides in the DNA methyltransferase catalytic domain. Additionally, all these algal sequences contain N-terminal extensions with conserved DNMT-RFD (DNA methyltransferase replication foci domain) and BAH domains (Figure 2.3), as observed in the canonical enzyme (Goll and Bestor, 2005; Ponger and Li, 2005).

Interestingly, *C. reinhardtii* contains two additional Dnmt1/MET1 related polypeptides, with high degree of homology to the catalytic domain but lacking conserved motifs in the N-terminal regions (Figure 2.3). One of these proteins, termed DMT1, has been characterized as a novel DNA methyltransferase with de novo nonselective cytosine methylation activity (Nishiyama *et al*., 2004). Moreover, it has been shown to localize to Chlamydomonas chloroplasts and to influence plastid DNA methylation and the uniparental inheritance of chloroplast genes (Nishiyama *et al*., 2004). Unexpectedly, in several microalgae, we also found putative DNA methyltransferases that cannot be clearly categorized (Table 1, Others). These predicted proteins contain DNA methyltransferase catalytic domains somewhat related to those of the Dnmt1/MET1

or the CMT subfamilies but they lack either N-terminal extensions or conserved domains in the N-terminal extensions. If correctly predicted and functional, some of these enzymes might be responsible for DNA methylation patterns or processes unique to microalgae (see below).

2.2.2 Biological role(s) of DNA cytosine methylation

Methylation of the fifth carbon of cytosine is a widespread modification present in bacteria, archaea, and eukaryotes (Feng *et al*., 2010; Zemach *et al*., 2010; Huff and Zilberman, 2014). However, methylation of cytosine residues has not been detected in several model organisms such as Saccharomyces cerevisiae, Schizosaccharomyces pombe, and C. elegans (Capuano *et al*., 2014), suggesting that this modification is not essential for eukaryotic life. Interestingly, some of the examined microalgae lack clearly identifiable homologs of Dnmt1/MET1, CMT, and Dnmt3/DRM methyltransferases (Table 1) but it is not certain that they are devoid of DNA cytosine methylation. For instance, the Prasinophyte algae (Ostreococcus lucimarinus, Bathycoccus prasinos, and M. pusilla CCMP1545) contain highly diverged DNA methyltransferases of the Dnmt5 subfamily and exhibit genomic CG methylation (Huff and Zilberman, 2014). In land plants or mammals, DNA cytosine methylation is often associated with the transcriptional silencing of transposable elements, repetitive DNA sequences, and some protein coding genes, and plays a critical role in genomic imprinting, X-chromosome inactivation, and chromosome stability (Feng *et al*., 2010; Law and Jacobsen, 2010). Gene body methylation is also highly conserved even though its precise function(s) remains an open question (Feng *et al*., 2010; Zemach *et al*., 2010; Huff and Zilberman, 2014). In contrast

to this wealth of information, the role(s) of DNA cytosine methylation in microalgae is still poorly understood.

Advances in sequencing technologies have enabled profiling of the genome methylation patterns of multiple species, including the microalgae *C. variabilis* NC64A, *V. carteri*, and *C. reinhardtii* (Feng *et al*., 2010; Zemach *et al*., 2010). In *C. variabilis*, genes are universally CG methylated, mostly within their bodies with a sharp drop at the promoters (Zemach *et al*., 2010). Interestingly, promoter methylation is inversely correlated with gene expression suggesting that, when it occurs, promoter-proximal methylation represses transcription. CHG methylation is also substantial but, similarly to land plants, concentrated in repetitive (presumably transposon) sequences and excluded from genes (Zemach *et al*., 2010). In contrast, the *V. carteri* genome is methylated to a much lower degree and exclusively in the CG context. Transposons and repeats are preferentially methylated but a weak negative relationship between promoter methylation and gene transcription was also observed (Zemach *et al*., 2010). In Volvox, DNA methylation had previously been implicated in the transcriptional silencing of introduced transgenes and a MET1-like methyltransferase was suggested to be involved in the maintenance of transgene and transposon methylation (Babinger *et al*., 2007).

In *C. reinhardtii*, DNA cytosine methylation was initially observed in chloroplast DNA during gametogenesis and implicated in the uniparental inheritance of mating type plus chloroplast DNA (Umen and Goodenough, 2001; Nishiyama *et al*., 2004). Recent findings indicated that the nuclear genome of this alga is also methylated at low levels and, similarly to Volvox, CG methylation is preferentially enriched in transposon sequences and to a much lower degree in gene bodies (Feng *et al*., 2010). Intriguingly,

CHG and CHH methylation was also observed uniformly along chromosomes and showed little enrichment in transposons or repeats (Feng *et al*., 2010). However, Chlamydomonas does not appear to contain CMT or Dnmt3/DRM homologs (Table 1 and Feng *et al*., 2010) and the methyltransferase(s) responsible for the later modifications has not been characterized. DNA methylation has also been implicated in the transcriptional silencing of introduced transgenes, particularly tandem repeats, in Chlamydomonas (Cerutti *et al*., 1997).

Collectively, the evidence suggests that, as in land plants and vertebrates, preferential DNA cytosine methylation of transposable elements and gene bodies has been conserved in some microalgae (Feng *et al*., 2010; Zemach *et al*., 2010). The methylation of transposons likely suppresses transcription and ensuing self-replication preserving genome integrity. Additionally, DNA methylation in promoters correlates negatively with gene expression in a subset of microalgae (Zemach *et al*., 2010) but it is not known whether certain protein genes are regulated by this modification. The pattern of genomic DNA methylation and the complement of DNA methyltransferases in species of the genus Chlorella, except for the lack of a Dnmt3/DRM homolog (Table 1), are similar to those in land plants. In contrast, C. reinhardtii and V. carteri methylate preferentially transposable elements, but predominantly in the CG (rather than the CHG/CHH) context and presumably by a somewhat different mechanism from that in land plants given the divergence in DNA methyltransferases (Table 1). Additionally, O. lucimarinus, M. pusilla, and B. prasinos, which possess DNA methyltransferases of the diverged Dnmt5 subfamily, exhibit densely clustered CG methylation in nucleosome linkers, possibly contributing to nucleosome positioning and proper chromatin

compaction in very small nuclei (Huff and Zilberman, 2014). Thus, the distribution and function(s) of DNA cytosine methylation in microalgae appear to be highly varied, making it difficult to extrapolate knowledge from well-characterized model systems as to its possible role(s) in the regulation of gene expression.

2.3 The RNA interference machinery in microalgae

2.3.1 Phylogenetic analysis and domain organization of core components of the RNAi machinery

Biochemical and genetic studies in multiple eukaryotes resulted in the identification of three key components of the RNAi machinery, namely Dicer, Argonaute-Piwi (AGO-Piwi), and RNA-dependent RNA Polymerase (RDR) (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Meister, 2013; Burroughs *et al*., 2014). In the best-characterized RNAi pathway, a variety of double stranded RNA (dsRNA) precursors are processed into small RNAs (sRNAs) by the RNaseIII-like endonuclease Dicer (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Rogers and Chen, 2013). These sRNAs are then incorporated into effector complexes, which include members of the AGO-Piwi family of proteins named after A. thaliana ARGONAUTE1 and Drosophila melanogaster Piwi (Cerutti and Casas-Mollano, 2006; Meister, 2013; Burroughs *et al*., 2014). AGO-Piwi proteins act as highly specialized small-RNA-binding modules. Based on structural studies, they are characterized by amino-terminal (N-terminal), PAZ (Piwi, Argonaute, Zwille), MID (middle) and PIWI domains, the latter being related to RNase H (Carthew and Sontheimer, 2009; Meister, 2013; Burroughs *et al*., 2014). Some AGO-Piwi polypeptides function as sRNA-guided

endonucleases that cleave complementary transcripts, whereas others lack endonucleolytic activity and may be part of complexes involved in non-degradative RNAi mechanisms such as translation repression (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Meister, 2013; Rogers and Chen, 2013). In certain species, RDRs also play an important role in RNAi, either by producing an initiating dsRNA trigger from singlestranded transcripts or by enhancing the RNAi response through amplification of the sRNA amount (Cerutti and Casas-Mollano, 2006; Burroughs *et al*., 2014).

The presence of Dicer, AGO-Piwi, and RDR polypeptides was examined in the 14 microalgae with sequenced genomes belonging to the Archaeplastida supergroup (Table 1). As previously reported (Cerutti and Casas-Mollano, 2006; Cerutti *et al*., 2011), core RNAi machinery components seem to be entirely absent from several algae with small nuclear genomes such as the red algae C. merolae and G. sulphuraria and the green algae O. lucimarinus, Ostreococcus tauri, B. prasinos, and M. pusilla CCMP1545 (Table 1). Micromonas sp. RCC299 appears to code for an AGO-Piwi related protein (Table 1). Yet, the corresponding gene is located in a region of the genome with no detectable homology to the closely related M. pusilla CCMP1545 (data not shown) and it is no clear if the encoded protein is actually functional. These observations are consistent with the hypothesis that the RNAi machinery appeared early during eukaryotic evolution and it has been lost independently in multiple lineages (Cerutti and Casas-Mollano, 2006; Cerutti *et al*., 2011). As a consequence, RNAi machinery components are present in algal species of each of the Glaucophyta, Rhodophyta, and Chlorophyta clades but with patchy distribution (Table 1).

AGO-Piwi proteins can often be identified in microalgae that also contain Dicer-like proteins (Table 1) and the domain organization of AGO-Piwi polypeptides has been

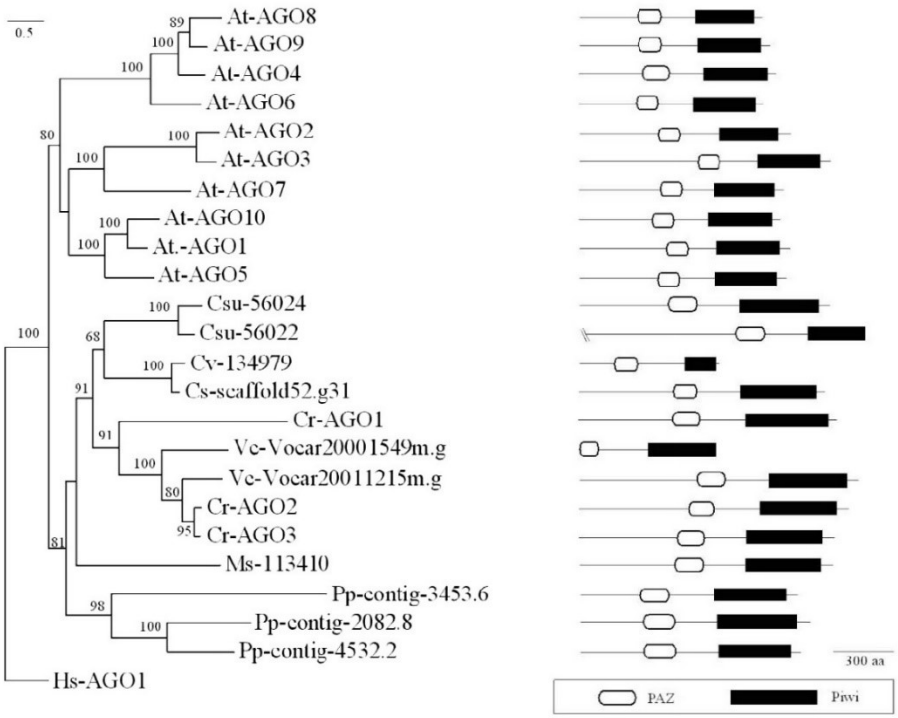

Figure 2.4. Maximum likelihood tree showing the phylogenetic relationship among Argonaute proteins. Sequences corresponding to AGO-Piwi proteins from different organisms were aligned using Muscle and the tree was drawn using the MEGA v6.06 program. Numbers indicate bootstrap values, >60%, based on 1000 pseudoreplicates. Species are designated by a two-letter or three-letter abbreviation preceding the name of each protein, as described in the legends to Figures 2.1 and 2.2, except for: Pp, Porphyridium purpureum. Accession numbers of proteins used to draw the tree are: At-AGO1, NP_849784; At-AGO2, NP_174413; At-AGO3, NP_174414; At-AGO4, NP_001189613; At-AGO5, NP_850110; At-AGO6, NP_180853; At-AGO7, NP_177103; At-AGO8, NP_197602; At-AGO9, NP_197613; At-AGO10, NP_199194; Cr-AGO1, Cr-AGO2, and Cr-AGO3, Cre02.g141050, Cre04.g214250, and Cre04.g689647 at http://phytozome.jgi.doe.gov; Cs-scaffold52.g31, from *Chlorella sorokiniana* draft genomic sequence assembly at the University of Nebraska-Lincoln; Csu-56022 and Csu-56024, PID 56022 and PID 56024 at http://phytozome.jgi.doe.gov; Cv-134979, PID 134979 at http://genome.jgi-psf.org/ ChlNC64A_1/ChlNC64A_1.home.html; Hs-AGO1, NP_03633; Ms-113410, PID 113410 at http://phytozome.jgi.doe.gov; Vc-Vocar20011215m.g and Vc-Vocar20001549m.g, Vocar20011215m and Vocar20001549 at http://phytozome.jgi.doe.gov; Pp-contig-4532.2, Pp-contig-2082.8, and Pp-contig-3453.6, evm.model.contig_4532.2, evm.model.contig_2082.8, and evm.model.contig_3453.6 at http://cyanophora.rutgers.edu. The SMART 7 database was

used to identify conserved protein motifs and the domain organization of the proteins is indicated on the right (see text for details).

well conserved in these species (Figure 2.4). In contrast, the multidomain structure of the Dicer enzymes of higher eukaryotes is not well maintained among protists (Cerutti and Casas-Mollano, 2006) and an RNaseIII motif might be all that is strictly required for function (Cerutti and Casas-Mollano, 2006; Patrick *et al*., 2009). RDR related proteins are structurally conserved but, consistent with an ancillary role in typical RNAi (Cerutti and Casas-Mollano, 2006), they appear to have a fairly limited taxonomic distribution among the examined microalgae (Table 1). Intriguingly, the glaucophyte C. paradoxa only seems to contain an RDR protein but its incomplete genome prevents a definitive assessment of the RNAi machinery components. Conversely, the red alga Porphyridium purpureum and, possibly, the green alga C. subellipsoidea encode all three core components of the RNAi machinery (Table 1), although the putative Coccomyxa Dicer is highly diverged (data not shown). In addition, AGO-Piwi encoding genes have undergone duplication in several microalgae (Figure 2.4), perhaps associated with diversification of functions as reported in land plants and metazoans (Casas-Mollano *et al*., 2008; Ghildiyal and Zamore, 2009; Meister, 2013).

In many eukaryotes, at least two major classes of small RNAs have been recognized: microRNAs (miRNAs) and small interfering RNAs (siRNAs) (Bartel, 2009; Ghildiyal and Zamore, 2009; Axtell, 2013; Rogers and Chen, 2013). MicroRNAs often modulate gene expression and originate from endogenous, single-stranded noncoding RNA transcripts or introns that fold into imperfect hairpin structures (Bartel, 2009; Ghildiyal and Zamore, 2009; Axtell, 2013; Rogers and Chen, 2013). siRNAs are

produced from long, near-perfect complementarity dsRNAs of diverse origins (including dsRNAs experimentally introduced into cells) and play various roles in post-transcriptional regulation of gene expression, suppression of viruses and transposable elements, DNA methylation, and/or heterochromatin formation (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009; Axtell, 2013; Rogers and Chen, 2013). The occurrence of small RNAs, potentially generated by Dicer-mediated processing, has been examined in very few of the sampled algal species.

C. reinhardtii has been shown to contain a complex set of endogenous sRNAs including miRNAs, phased siRNAs, as well as siRNAs originating from transposons and repeated DNA regions (Molnar *et al*., 2007; Zhao *et al*., 2007; Yamasaki *et al*., 2013; Voshall *et al*., 2014). Likewise, V. carteri has been demonstrated to possess miRNAs, many of them being preferentially enriched in either gonidia or somatic cells (Li *et al*., 2014). Interestingly, there appears to be little conservation of miRNA genes among algal species or with those encoded in the genomes of land plants or animals (Cerutti *et al*., 2011; Li *et al*., 2014; Voshall *et al*., 2014).

2.3.2 Biological role(s) of RNA interference

The presence of core components of the RNAi machinery and of endogenous miRNAs/siRNAs in certain microalgae suggest that RNA-mediated silencing may play a biological role(s) in these organisms. However, it is also apparent that the RNAi machinery seems to be dispensable for cellular life since it has been entirely lost or extensively modified in several species. In a wide range of eukaryotes, sRNA-mediated mechanisms have been implicated in multiple processes including transposon silencing,

resistance to viruses, regulation of endogenous gene expression, heterochromatin formation, DNA methylation, DNA repair, and maintenance of genome stability (Cerutti and Casas-Mollano, 2006; Carthew and Sontheimer, 2009; Ohsawa *et al*., 2013; Rogers and Chen, 2013; Oliver *et al*., 2014; Zhong *et al*., 2014). In contrast, very little is known about the actual function(s) of RNA-mediated silencing in microalgae. C. reinhardtii has undergone duplication of key RNAi components (Casas-Mollano *et al*., 2008) and contains three Dicers and three Argonautes (Table 1). One of the Dicer proteins, DCL1, appears to be part of a siRNA pathway that has specialized for the control of transposable elements (Casas-Mollano *et al*., 2008). However, C. reinhardtii also relies on a DCL1-independent, transcriptional silencing mechanism(s) for transposon repression (Casas-Mollano *et al*., 2007; Shaver *et al*., 2010). Interestingly, this chromatin-mediated silencing is sensitive to temperature, being much more effective at 17 ℃ than at 25 ℃ (Cerutti *et al*., 1997). Conversely, RNAi dependent post-transcriptional gene repression has been shown to be more efficient at 25–29 ℃ in both invertebrates and land plants (Fortier and Belote, 2000; Szittya *et al*., 2003). Thus, it is tempting to speculate that in C. reinhardtii, multiple, partly independent silencing mechanisms may operate to suppress reliably transposon mobilization over a wide range of environmental conditions (Casas-Mollano *et al*., 2008; Cerutti *et al*., 2011). In C. variabilis NC64A recent transcriptome analyses revealed that components of certain RNA-silencing pathways are up regulated after infection with the PBCV-1 virus (Rowe *et al*., 2013). Thus, the limited available evidence does suggest that RNAi is involved in defense responses against transposable elements and, possibly, viruses in some microalgae.

In higher eukaryotes, RNAi also plays an important role in the regulation of endogenous gene expression through miRNAs and other small RNAs (Bartel, 2009; Ghildiyal and Zamore, 2009; Axtell, 2013; Rogers and Chen, 2013). Within the examined microalgae, candidate miRNAs have been experimentally identified in *C. reinhardtii* (Molnar *et al*., 2007; Zhao *et al*., 2007; Yamasaki *et al*., 2013; Voshall *et al*., 2014) and in *V. carteri* (Li *et al*., 2014), and potential endogenous targets have been predicted by computational approaches (Molnar *et al*., 2007; Zhao *et al*., 2007; Li *et al*., 2014; Voshall *et al*., 2014). However, target gene identification is a challenging problem (Thomas *et al*., 2010; Voshall *et al*., 2014) and the false-positive prediction rate for miRNA targets in microalgae is currently unknown. Interestingly, most characterized land plant miRNAs appear to regulate transcripts with highly complementary binding sites and preferentially trigger their endonucleolytic cleavage by Argonaute proteins (Bartel, 2009; Axtell, 2013; Rogers and Chen, 2013). This also appears to be the case for some miRNAs in both *C. reinhardtii* and *V. carteri*, where expected target RNA cleavage products have been experimentally detected (Molnar *et al*., 2007; Zhao *et al*., 2007; Li *et al*., 2014; Voshall *et al*., 2014). Yet, very few of the reported Chlamydomonas or Volvox miRNAs have identifiable targets with near perfect complementarity (Molnar *et al*., 2007; Li *et al*., 2014; Voshall *et al*., 2014) and recent evidence suggests that miRNA regulation of transcript expression in Chlamydomonas may operate, at least for certain targets, by translation repression (Ma *et al*., 2013; Yamasaki *et al*., 2013; Voshall *et al*., 2014).

In summary, the RNA interference mechanism appears to be entirely absent from some microalgae and its biological role(s) in the species that possess core RNAi machinery components is poorly understood. Limited evidence suggests that a siRNA

pathway may operate as a defense mechanism against transposon mobilization and, possibly, in antiviral immunity. A miRNA pathway, when present, may contribute to endogenous gene regulation. However, the identification of genuine miRNA targets remains challenging and, to date, no specific metabolic or physiological process controlled or modulated by miRNAs has been clearly defined in microalgae. Other possible functions of RNAi in phenomena such as heterochromatin formation, DNA methylation, DNA repair, or maintenance of genome stability, to our knowledge, have not been explored in these aquatic organisms.

Gene silencing mechanisms and biofuel/biomaterial production

Algae exhibit the potential for manufacturing a wide range of biofuel precursors and bioproducts, and advances in algal genomics, genetic engineering and synthetic biology may facilitate the development of industrial strains suitable to specific production purposes. However, a serious limitation to strain improvement is our incomplete understanding of gene and metabolic network regulation in most algal species. In many eukaryotes, gene silencing mechanisms have been implicated in context-dependent reversible gene repression, a critical component of gene networks involved in stress responses and in developmental programs (Law and Jacobsen, 2010; Bannister and Kouzarides, 2011; Ohsawa *et al*., 2013; Derkacheva and Hennig, 2014). Interestingly, most microalgae accumulate biofuel-precursor compounds, such as starch and triacylglycerols (TAGs), primarily under stress conditions (Hu *et al*., 2008; Radakovits *et al*., 2010; La Russa *et al*., 2012; Liu and Benning, 2013). Transcriptome analyses in C. reinhardtii (reviewed by Liu and Benning, 2013) have revealed that a subset of genes

involved in TAG biosynthesis, such as those encoding certain acyl-CoA: diacylglycerol acyltransferases, display greatly enhanced or almost exclusive expression in nitrogen-deprived cells. These observations suggest that gene silencing mechanisms may indeed be involved in modulating metabolic responses to stress in microalgae. Specific genes for storage compound biosynthesis appear to be repressed under normal environmental conditions, when they may not be needed, and activated primarily under a variety of stresses. Yet, to our knowledge, direct evidence for a role(s) of epigenetic silencing mechanisms in microalgal gene regulation is currently missing.

Notwithstanding its endogenous role, RNA interference also provides a tool for functional genomic analyses in microalgae (reviewed by Cerutti *et al*., 2011). In some species, transient gene repression has been achieved by introduction into algal cells of non-integrative dsRNA/sRNA producing plasmids or of exogenously synthesized dsRNAs/sRNAs (Cerutti *et al*., 2011; Gimpel *et al*., 2013). In microalgae with available genome transformation methodologies stable and heritable RNAi has been developed, relying generally on the production of hairpin dsRNAs or artificial microRNAs from genome-integrated transgenes (Radakovits *et al*., 2010; Cerutti *et al*., 2011; Gimpel *et al*., 2013; Liu and Benning, 2013). These approaches have been used to suppress endogenous gene expression in microalgae, in order to characterize functionally genes involved in lipid metabolism (Gimpel *et al*., 2013; Liu and Benning, 2013) as well as to engineer algal strains with altered traits for biotechnological purposes, such as optimizing antenna size to improve light utilization and photosynthetic efficiency (Gimpel *et al*., 2013). RNAi technology (Cerutti *et al*., 2011), in conjunction with system level 'omics' approaches, may contribute much-needed insight into gene function, metabolic pathways,

and regulatory networks in microalgae, as well as provide a valuable tool for the genetic engineering of enhanced traits.

Conclusion and perspective

Microalgae exhibit the potential for manufacturing biofuel precursors and a range of valuable bioproducts. However, a serious limitation to rational strain improvement is our incomplete understanding of gene function and regulation in most species. Gene silencing mechanisms appear to be widespread in the Archaeplastida microalgae, including post-translational histone modifications, DNA methylation, and small RNA directed pathways. Yet, there also appears to be great diversity in the silencing systems present in individual algal species. Future challenges will involve defining whether these epigenetic processes regulate storage compound metabolic networks and developing improved tools for the genetic engineering of production strain.

References

Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., Shilatifard, A., Workman, J., Zhang, Y., 2007. New nomenclature for chromatin-modifying enzymes. Cell 131, 633–636.

Arabidopsis Genome Initiative, 2000. Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408, 796–815.

Axtell, M.J., 2013. Classification and comparison of small RNAs from plants. Annu.

Rev. Plant Biol. 64, 137–159.

Babinger, P., Völkl, R., Cakstina, I., Maftei, A., Schmitt, R., 2007. Maintenance DNA methyltransferase (Met1) and silencing of CpG-methylated foreign DNA in Volvox carteri. Plant Mol. Biol. 63, 325–336.

Bannister, A.J., Kouzarides, T., 2011. Regulation of chromatin by histone modifications. Cell Res. 21, 381–395.

Bartel, D.P., 2009. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233.

Bhattacharya, D., Price, D.C., Chan, C.X., Qiu, H., Rose, N., Ball, S., Weber, A.P., Arias, M.C., Henrissat, B., Coutinho, P.M., Krishnan, A., Zäuner, S., Morath, S., Hilliou, F., Egizi, A., Perrineau, M.M., Yoon, H.S., 2013. Genome of the red alga Porphyridium purpureum. Nat. Commun. 4, 1941.

Blanc, G., Duncan, G., Agarkova, I., Borodovsky, M., Gurnon, J., Kuo, A., Lindquist, E., Lucas, S., Pangilinan, J., Polle, J., Salamov, A., Terry, A., Yamada, T., Dunigan, D.D., Grigoriev, Claverie, J.M., Van Etten, J.L., 2010. The Chlorella variabilis NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. Plant Cell 22, 2943–2955.

Blanc, G., Agarkova, I., Grimwood, J., Kuo, A., Brueggeman, A., Dunigan, D.D., Gurnon, J., Ladunga, I., Lindquist, E., Lucas, S., Pangilinan, J., Pröschold, T., Salamov, A., Schmutz, J., Weeks, D., Yamada, T., Lomsadze, A., Borodovsky, M., Claverie, J.M., Grigoriev, I.V., Van Etten, J.L., 2012. The genome of the polar eukaryotic microalga Coccomyxa subellipsoidea reveals traits of cold adaptation. Genome Biol. 13 (5), R39.

Burroughs, A.M., Ando, Y., Aravind, L., 2014. New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. Wiley Interdiscip. Rev. RNA 5, 141–181.

Capuano, F., Mulleder, M., Kok, R., Blom, H.J., Ralser, M., 2014. Cytosine DNA methylation is found in Drosophila melanogaster but absent in Saccharomyces cerevisiae, Schizosaccharomyces pombe, and other yeast species. Anal. Chem. 86, 3697–3702.

Carthew, R.W., Sontheimer, E.J., 2009. Origins and mechanisms of miRNAs and siRNAs. Cell 136, 642–655.

Casas-Mollano, J.A., van Dijk, K., Eisenhart, J., Cerutti, H., 2007. SET3p monomethylates histone H3 on lysine 9 and is required for the silencing of tandemly repeated transgenes in Chlamydomonas. Nucleic Acids Res. 35, 939–950.

Casas-Mollano, J.A., Rohr, J., Kim, E.J., Balassa, E., van Dijk, K., Cerutti, H., 2008. Diversification of the core RNA interference machinery in Chlamydomonas reinhardtii and the role of DCL1 in transposon silencing. Genetics 179, 69–81.

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

Cerutti, H., Casas-Mollano, J.A., 2006. On the origin and functions of RNA-mediated silencing: from protists to man. Curr. Genet. 50, 81–99.

Cerutti, H., Ma, X., Msanne, J., Repas, T., 2011. RNA-mediated silencing in algae: biological roles and tools for the analysis of gene function. Eukaryot. Cell 10, 1164–1172.

Chisti, Y., 2013. Constraints to commercialization of algal fuels. J. Biotechnol. 167, 201–214.

Derelle, E., Ferraz, C., Rombauts, S., Rouzé, P., Worden, A.Z., Robbens, S., Partensky, F., Degroeve, S., Echeynié, S., Cooke, R., Saeys, Y., Wuyts, J., Jabbari, K., Bowler, C., Panaud, O., Piégu, B., Ball, S.G., Ral, J.P., Bouget, F.Y., Piganeau, G., De Baets, B., Picard, A., Delseny, M., Demaille, J., Van de Peer, Y., Moreau, H., 2006. Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features. Proc. Natl. Acad. Sci. U.S.A. 103, 11647–11652.

Derkacheva, M., Hennig, L., 2014. Variations on a theme: polycomb group proteins in plants. J. Exp. Bot. 65, 2769–2784.

Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., Hetzel, J., Wohlschlegel, J.A., Pradhan, S., Patel, D.J., Jacobsen, S.E., 2012. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell 151, 167–180.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., Ukomadu, C., Sadler, K.C., Pradhan, S., Pellegrini, M., Jacobsen, S.E., 2010. Conservation and divergence of methylation patterning in plants and animals. Proc. Natl. Acad. Sci. U.S.A. 107, 8689–8694.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Fortier, E., Belote, J.M., 2000. Temperature-dependent gene silencing by an expressed inverted repeat in Drosophila. Genesis 26, 240–244.

Ghildiyal, M., Zamore, P.D., 2009. Small silencing RNAs: an expanding universe. Nat. Rev. Genet. 10, 94–108.

Gimpel, J.A., Specht, E.A., Georgianna, D.R., Mayfield, S.P., 2013. Advances in microalgae engineering and synthetic biology applications for biofuel production. Curr. Opin. Chem. Biol. 17, 489–495.

Goll, M.G., Bestor, T.H., 2005. Eukaryotic cytosine methyltransferases. Annu. Rev. Biochem. 74, 481–514.

Griffiths, M.J., Harrison, S.T.L., 2009. Lipid productivity as a key characteristic for choosing algal species for biodiesel production. J. Appl. Phycol. 21, 493–507.

Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., Darzins, A., 2008. Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. Plant J. 54, 621–639.

Huang, Y., Liu, C., Shen, W.-H., Ruan, Y., 2011. Phylogenetic analysis and classification of the Brassica rapa SET-domain protein family. BMC Plant Biol. 11, 175.

Huff, J.T., Zilberman, D., 2014. Dnmt1-independent CG methylation contributes to nucleosome positioning in diverse eukaryotes. Cell 156, 1286–1297.

Krauss, V., 2008. Glimpses of evolution: heterochromatic histone H3K9 methyltransferase left its mark behind. Genetica 133, 93–106.

Krishnan, S., Horowitz, S., Trievel, R.C., 2011. Structure and function of histone H3 lysine 9 methyltransferases and demethylases. ChemBioChem 12, 254–263.

La Russa, M., Bogen, C., Uhmeyer, A., Doebbe, A., Filippone, E., Kruse, O., Mussgnug,

J.H., 2012. Functional analysis of three type-2 DGAT homologue genes for triacylglycerol production in the green microalga Chlamydomonas reinhardtii. J. Biotechnol. 162, 13–20.

Law, J.A., Jacobsen, S.E., 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. 11, 204–220. Lee, D.H., 2011. Algal biodiesel economy and competition among biofuels. Bioresour. Technol. 102, 43–49.

Leite, G.B., Abdelaziz, A.E., Hallenbeck, P.C., 2013. Algal biofuels: challenges and opportunities. Bioresour. Technol. 145, 134–141.

Li, J., Wu, Y., Qi, Y., 2014. MicroRNAs in a multicellular green alga Volvox carteri. Sci. China Life Sci. 57, 36–45.

Liu, B., Benning, C., 2013. Lipid metabolism in microalgae distinguishes itself. Curr. Opin. Biotechnol. 24, 300–309.

Ma, X., Kim, E.J., Kook, I., Ma, F., Voshall, A., Moriyama, E., Cerutti, H., 2013. Small interfering RNA-mediated translation repression alters ribosome sensitivity to inhibition by cycloheximide in Chlamydomonas reinhardtii. Plant Cell 25, 985-998.

Matsuzaki, M., Misumi, O., Shin-I, T., Maruyama, S., Takahara, M., Miyagishima, S.Y., Mori, T., Nishida, K., Yagisawa, F., Nishida, K., Yoshida, Y., Nishimura, Y., Nakao, S., Kobayashi, T., Momoyama, Y., Higashiyama, T., Minoda, A., Sano, M., Nomoto, H., Oishi, K., Hayashi, H., Ohta, F., Nishizaka, S., Haga, S., Miura, S., Morishita, T., Kabeya, Y., Terasawa, K., Suzuki, Y., Ishii, Y., Asakawa, S., Takano, H., Ohta, N., Kuroiwa, H., Tanaka, K., Shimizu, N., Sugano, S., Sato, N.,

Nozaki, H., Ogasawara, N., Kohara, Y., Kuroiwa, T., 2004. Genome sequence of the ultrasmall unicellular red alga Cyanidioschyzon merolae 10D. Nature 428, 653–657.

Meister, G., 2013. Argonaute proteins: functional insights and emerging roles. Nat. Rev. Genet. 14, 447–459.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., Marshall, W.F., Qu, L.H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y.,

Martínez, D., Ngau, W.C., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., Grossman, A.R., 2007. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318, 245–251.

Molnar, A., Schwach, F., Studholme, D.J., Thuenemann, E.C., Baulcombe, D.C., 2007. MiRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447, 1126–1129.

Moreau, H., Verhelst, B., Couloux, A., Derelle, E., Rombauts, S., Grimsley, N., Van Bel, M., Poulain, J., Katinka, M., Hohmann-Marriott, M.F., Piganeau, G., Rouzé, P., Da Silva, C., Wincker, P., Van de Peer, Y., Vandepoele, K., 2012. Gene functionalities and genome structure in Bathycoccus prasinos reflect cellular specializations at the base of the green lineage. Genome Biol. 13, R74.

Mujtaba, S., Manzur, K.L., Gurnon, J.R., Kang, M., Van Etten, J.L., Zhou, M.M., 2008. Epigenetic transcriptional repression of cellular genes by a viral SET protein. Nat. Cell Biol. 10, 1114–1122.

Nishiyama, R., Wada, Y., Mibu, M., Yamaguchi, Y., Shimogawara, K., Sano, H., 2004. Role of a nonselective de novo DNA methyltransferase in maternal inheritance of chloroplast genes in the green alga Chlamydomonas reinhardtii. Genetics 168, 809–816.

Ohsawa, R., Seol, J.H., Tyler, J.K., 2013. At the intersection of non-coding transcription, DNA repair, chromatin structure, and cellular senescence. Front. Genet. 4, 136.

Oliver, C., Santos, J.L., Pradillo, M., 2014. On the role of some ARGONAUTE proteins in meiosis and DNA repair in Arabidopsis thaliana. Front. Plant Sci. 5, 177.

Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., Dupont, C., Jorgensen, R., Derelle, E., Rombauts, S., Zhou, K., Otillar, R., Merchant, S.S., Podell, S., Gaasterland, T., Napoli, C., Gendler, K., Manuell, A., Tai, V., Vallon, O., Piganeau, G., Jancek, S., Heijde, M., Jabbari, K., Bowler, C., Lohr, M., Robbens, S., Werner, G., Dubchak, I., Pazour, G.J., Ren, Q., Paulsen, I., Delwiche, C., Schmutz, J., Rokhsar, D., Van de Peer, Y., Moreau, H., Grigoriev, I.V., 2007. The tiny eukaryote Ostreococcus provides genomic insights into the paradox of plankton speciation. Proc. Natl. Acad. Sci. U.S.A. 104, 7705–7710.

Patrick, K.L., Shi, H., Kolev, N.G., Ersfeld, K., Tschudi, C., Ullu, E., 2009. Distinct and overlapping roles of two Dicer-like proteins in the RNA interference pathways of the ancient eukaryote Trypanosoma brucei. Proc. Natl. Acad. Sci. U.S.A. 106, 17933–17938.

Ponger, L., Li, W.-H., 2005. Evolutionary diversification of DNA methyltransferases in eukaryotic genomes. Mol. Biol. Evol. 22, 1119–1128.

Price, D.C., Chan, C.X., Yoon, H.S., Yang, E.C., Qiu, H., Weber, A.P., Schwacke, R., Gross, J., Blouin, N.A., Lane, C., Reyes-Prieto, A., Durnford, D.G., Neilson, J.A., Lang, B.F., Burger, G., Steiner, J.M., Löffelhardt, W., Meuser, J.E., Posewitz, M.C., Ball, S., Arias, M.C., Henrissat, B., Coutinho, P.M., Rensing, S.A., Symeonidi, A., Doddapaneni, H., Green, B.R., Rajah, V.D., Boore, J., Bhattacharya, D., 2012.

Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. Science 335, 843–847.

Prochnik, S.E., Umen, J., Nedelcu, A.M., Hallmann, A., Miller, S.M., Nishii, I., Ferris, P., Kuo, A., Mitros, T., Fritz-Laylin, L.K., Hellsten, U., Chapman, J., Simakov, O., Rensing, S.A., Terry, A., Pangilinan, J., Kapitonov, V., Jurka, J., Salamov, A., Shapiro, H., Schmutz, J., Grimwood, J., Lindquist, E., Lucas, S., Grigoriev, I.V., Schmitt, R., Kirk, D., Rokhsar, D.S., 2010. Genomic analysis of organismal complexity in the multicellular green alga Volvox carteri. Science 329, 223–226.

Radakovits, R., Jinkerson, R.E., Darzins, A., Posewitz, M.C., 2010. Genetic engineering of algae for biofuel production. Eukaryot. Cell 9, 486–501.

Rajakumara, E., Law, J.A., Simanshu, D.K., Voigt, P., Johnson, L.M., Reinberg, D., Patel, D.J., Jacobsen, S.E., 2011. A dual flip-out mechanism for 5mC recognition by the Arabidopsis SUVH5 SRA domain and its impact on DNA methylation and H3K9 dimethylation in vivo. Genes Dev. 25, 137–152.

Rogers, K., Chen, X., 2013. Biogenesis, turnover, and mode of action of plant miRNAs. Plant Cell 25, 2383–2399.

Rowe, J.M., Dunigan, D.D., Blanc, G., Gurnon, J.R., Xia, Y., Van Etten, J.L., 2013. Evaluation of higher plant virus resistance genes in the green alga, Chlorella variabilis NC64A, during the early phase of infection with Paramecium bursaria chlorella virus-1. Virology 442, 101–113.

Saze, H., Kakutani, T., 2011. Differentiation of epigenetic modifications between transposons and genes. Curr. Opin. Plant Biol. 14, 81–87.

Schönknecht, G., Chen, W.H., Ternes, C.M., Barbier, G.G., Shrestha, R.P., Stanke, M., Bräutigam, A., Baker, B.J., Banfield, J.F., Garavito, R.M., Carr, K., Wilkerson, C., Rensing, S.A., Gagneul, D., Dickenson, N.E., Oesterhelt, C., Lercher, M.J.,

Weber, A.P., 2013. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 339, 1207–1210.

Shaver, S., Casas-Mollano, J.A., Cerny, R.L., Cerutti, H., 2010. Origin of the polycomb repressive complex 2 and gene silencing by an E(z) homolog in the unicellular alga Chlamydomonas. Epigenetics 5, 301–312.

Strenkert, D., Schmollinger, S., Schroda, M., 2013. Heat shock factor 1 counteracts epigenetic silencing of nuclear transgenes in Chlamydomonas reinhardtii. Nucleic Acid Res. 41, 5273–5289.

Szittya, G., Silhavy, D., Molnár, A., Havelda, Z., Lovas, A., Lakatos, L., Bánfalvi, Z., Burgyán, J., 2003. Low temperature inhibits RNA silencing-mediated defence by the control of siRNA generation. EMBO J. 22, 633–640.

Thomas, M., Lieberman, J., Lai, A., 2010. Desperately seeking microRNA targets. Nat. Struct. Mol. Biol. 17, 1169–1174.

Tirichine, L., Bowler, C., 2011. Decoding algal genomes: tracing back the history of photosynthetic life on Earth. Plant J. 66, 45–57.

Umen, J.G., Goodenough, U.W., 2001. Chloroplast DNA methylation and inheritance in Chlamydomonas. Genes Dev. 15, 2585–2597.

Voshall, A., Kim, E.J., Ma, X., Moriyama, E.N., Cerutti, H., 2014. Identification of AGO3 associated miRNAs and computational prediction of their targets in the green alga Chlamydomonas reinhardtii. Genetics (submitted for publication).

Wei, H., Zhou, M.M., 2010. Viral-encoded enzymes that target host chromatin functions. Biochim. Biophys. Acta 1799, 296–301.

Worden, A.Z., Lee, J.H., Mock, T., Rouzé, P., Simmons, M.P., Aerts, A.L., Allen, A.E.,

Cuvelier, M.L., Derelle, E., Everett, M.V., Foulon, E., Grimwood, J., Gundlach, H., Henrissat, B., Napoli, C., McDonald, S.M., Parker, M.S., Rombauts, S., Salamov, A., Von Dassow, P., Badger, J.H., Coutinho, P.M., Demir, E., Dubchak, I., Gentemann, C., Eikrem, W., Gready, J.E., John, U., Lanier, W., Lindquist, E.A., Lucas, S., Mayer, K.F., Moreau, H., Not, F., Otillar, R., Panaud, O., Pangilinan, J., Paulsen, I., Piegu, B., Poliakov, A., Robbens, S., Schmutz, J., Toulza, E., Wyss, T., Zelensky, A., Zhou, K., Armbrust, E.V., Bhattacharya, D., Goodenough, U.W., Van de Peer, Y., Grigoriev, I.V., 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes micromonas. Science 324, 268–272.

Worden, A.Z., Allen, A.E., 2010. The voyage of the microbial eukaryote. Curr. Opin. Microbiol. 13, 652–660.

Yamasaki, T., Voshall, A., Kim, E.J., Moriyama, E., Cerutti, H., Ohama, T., 2013. Complementarity to a miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga Chlamydomonas reinhardtii. Plant J. 76, 1045–1056.

Zemach, A., McDaniel, I.E., Silva, P., Zilberman, D., 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science 328, 916–919.

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G.J., Wang, X.J., Qi, Y., 2007. A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes Dev. 21, 1190–1203.

Zheng, S., Wang, J., Feng, Y., Wang, J., Ye, K., 2012. Solution structure of MSL2 CXC domain reveals an unusual Zn3Cys9 cluster and similarity to Pre-SET domains of

histone lysine methyltransferases. PLoS One 7, e45437.

Zhong, X., Du, J., Hale, C.J., Gallego-Bartolome, J., Feng, S., Vashisht, A.A., Chory, J., Wohlschlegel, J.A., Patel, D.J., Jacobsen, S.E., 2014. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. Cell 157, 1050–1060.

CHAPTER 3

MIRNAS IN THE ALGA CHLAMYDOMONAS REINHARDTII ARE NOT PHYLOGENETICALLY CONSERVED AND PLAY A LIMITED ROLE IN RESPONSES TO NUTRIENT DEPRIVATION

Abstract

The unicellular alga *Chlamydomonas reinhardtii* contains many types of small RNAs (sRNAs) but the biological role(s) of bona fide microRNAs (miRNAs) remains unclear. To address their possible function(s) in responses to nutrient availability, we examined miRNA expression in cells cultured under different trophic conditions (mixotrophic in the presence of acetate or photoautotrophic in the presence or absence of nitrogen). We also reanalyzed miRNA expression data in Chlamydomonas subject to sulfur or phosphate deprivation. Several miRNAs were differentially expressed under the various trophic conditions. However, in transcriptome analyses, the majority of their predicted targets did not show expected changes in transcript abundance, suggesting that they are not subject to miRNA-mediated RNA degradation. Mutant strains, defective in sRNAs or in ARGONAUTE3 (a key component of sRNA- mediated gene silencing), did not display major phenotypic defects when grown under multiple nutritional regimes. Additionally, Chlamydomonas miRNAs were not conserved, even in algae of the closely related Volvocaceae family, and many showed features resembling those of recently evolved, species-specific miRNAs in the genus Arabidopsis. Our results suggest that, in

C. reinhardtii, miRNAs might be subject to relatively fast evolution and have only a minor, largely modulatory role in gene regulation under diverse trophic states.

Introduction

MicroRNAs are short RNA molecules (~20–24 nt in length) that generally function as negative regulators of gene expression, by binding complementary sequences in target transcripts and leading to translation repression and/or mRNA degradation[1–6]. In higher eukaryotes, miRNAs have been implicated in the control of many biological processes such as development, metabolism or stress responses[1–3, 7–9]. In contrast, the biological role(s) of miRNAs in unicellular organisms such as the alga Chlamydomonas reinhardtii[4, 5] remains unclear, although they have been recently proposed to modulate adaptation to abiotic stress[10]. In land plants, a subset of miRNAs is differentially or uniquely expressed under nutrient deprivation[8, 9, 11–13]. However, these condition-specific miRNAs often have very few, if any, targets within the pathways directly involved in the response to nutrient limitation, making it difficult to assess their overall significance[8, 9, 11–13].

Sulfur deprivation has been reported to induce differential miRNA expression in C. reinhardtii but, similarly to observations in higher plants, very few predicted targets appeared to be involved in mechanisms responding directly to sulfur deficiency[6]. In addition, it is often difficult to identify genuine miRNA-regulated transcripts in Chlamydomonas[14]. For instance, several putative targets showed mRNA up-regulation when the miRNAs predicted to target them also increased in abundance[6, 15]. This

expression pattern makes it unlikely that the predicted targets, if genuine, are regulated via miRNA-mediated transcript degradation although they could still be translationally modulated[14, 16, 17]. A further challenge to characterize miRNA function in *C. reinhardtii* is posed by the apparent lack of conservation among algal miRNAs and those identified in higher plants and animals[14, 18–20].

Plasticity in miRNA populations has been reported in both animal and plant lineages, as reflected by the number of miRNAs that are either species-specific or limited to closely related species[2, 21–26]. Within the genus Arabidopsis, there is evidence that miRNA genes may arise from inverted duplication of sequences or from spontaneous mutations in sequences capable of forming hairpin structures[2, 20, 23, 24]. This generation of novel miRNA genes from random sequences may account for the large number of miRNAs corresponding uniquely to *A. thaliana* or *A. lyrata* (~13% of their total miRNA populations) despite their relatively recent evolutionary divergence[24, 26]. These species-specific miRNAs appear to be lowly expressed, whereas miRNAs conserved among plant species generally show higher expression levels[25, 26]. Furthermore, many novel miRNAs have no experimentally confirmed targets and their function(s) is largely unknown[24-26] although some could conceivably play a role in lineage-specific processes such as responses to environmental conditions unique to the habitat of individual species[7, 9, 12, 27]. In order to characterize the possible roles of miRNAs in *C. reinhardtii*, particularly in response to nutrient availability, we investigated changes in miRNA populations and their putative targets in Chlamydomonas cells grown under mixotrophic conditions (in the presence of acetate) and under photoautotrophic conditions with or without a source of nitrogen. We also examined the phenotypes of mutant strains, lacking sRNAs or

defective in a core component of the RNA interference (RNAi) machinery, under nutrient deprived conditions. Subsets of differentially expressed miRNAs were identified under the various trophic conditions but very few of their predicted targets displayed expected changes in transcript abundance (assuming regulation by miRNA-triggered RNA degradation) or coded for proteins involved in direct responses to nutrient deficiency. The identified Chlamydomonas miRNAs were not conserved even in related green algal lineages and many showed relatively low expression levels, similarly to the recently evolved miRNAs characterized in higher plants. Interestingly, the RNAi defective mutants displayed slight phenotypic defects, suggesting that miRNAs might not play an essential role in endogenous gene regulation under the conditions examined.

Results

Changes in C. reinhardtii miRNA populations under various trophic conditions.

Small RNAs associated with ARGONAUTE3 (AGO3), a key component of the RNAi machinery in C. reinhardtii[14, 28], were isolated by co-immunoprecipitation with FLAG-tagged AGO3, sRNA libraries constructed and then analyzed by deep sequencing (see Methods). From these sequences, miRNAs were predicted based on the criteria outlined by Tarver *et al*.[29]. To differentiate miRNAs from other sRNAs, all genome mapped reads (see Methods) were clustered by genomic location such that within each cluster adjacent reads were no more than 200 nt apart, regardless of strand[14]. Genomic sequences for each strand of each cluster were then folded using RNAfold to determine

their secondary structure[14]. In order to be classified as a miRNA precursor, a cluster was required to fold into a hairpin and have no more than two predominant 5' processing sites[14, 29]. In addition, the main reads (i.e., the greatest abundance reads in each cluster; usually representing >90% of the locally mapped reads) were required to have no more than four mismatches in the complementary arm of the hairpin[14, 29]. By using these criteria, we identified 120 candidate miRNAs, across three growth conditions, co-immunoprecipitating with FLAG-tagged AGO3 (Figures 3.1A and S1). These sequences included the 45 miRNAs previously identified in cells grown under mixotrophic conditions in TAP (Tris-Acetate-Phosphate) medium[14] as well as 75 additional candidate miRNAs (Table S1A). Most miRNAs (83 of 120) were detected in photoautotrophically grown cells in nutrient replete high salt medium (HS + N), but only 14 were identified primarily in these cells whereas the majority (69 of 83) was shared between at least two trophic conditions (Figure 3.1A and Table S1A). On the other hand, Chlamydomonas grown photoautotrophically in nitrogen deprived medium (HS − N) had 20 condition-prevalent miRNAs (Figure 3.1A).

The population of AGO3-associated miRNAs clearly varies among cells grown under different trophic conditions (Table S1A). However, virtually none of the mature miRNA sequences is completely missing from libraries from any condition (Figure S1 and Table S1A). Most precursor miRNAs appear to be transcribed and processed under all growth conditions but the generated sRNA sequences may only meet the criteria to be classified as miRNAs under one or two of the examined nutritional regimes. Often, this occurs because the mature miRNA sequence may not represent at least 90% of the local reads matching to the precursor miRNA hairpin (Figure S1), one of the criteria for

prediction of canonical miRNAs[14, 29]. When considering read abundance, most miRNAs identified as such primarily in cells grown under a specific trophic condition (i.e., the 16 miRNAs in TAP, the 14 miRNAs in HS + N and the 20 miRNAs in HS − N) were also present at their highest levels in the libraries from that same condition (Table S1A). Yet, there were also some inconsistencies. For instance, miR_t35 and miR_t79 were classified as miRNAs in libraries from cells grown in HS − N but they seemed to be more prevalent in cells cultured in TAP (Table S1A). The abundance of AGO3-associated miRNAs identified in Chlamydomonas grown under multiple trophic conditions (i.e., the 32 miRNAs common to all three conditions examined, Figure 3.1A) remained relatively constant or differed depending on the nutritional regime (Table S1A). The expression of a subset of these miRNAs was validated by northern blot analyses of the Maa7-IR44s strain (containing the FLAG-tagged AGO3 protein), the parental strain CC-124, and a previously described mutant strain, Mut-2014, virtually devoid of small RNAs (Figure 3A). The U6 snRNA, whose abundance remains fairly stable under the examined conditions (Figure S2), was used as a loading control. Even though the RNA blots measure total cellular miRNA abundance whereas the libraries reflect AGO3-associated miRNA abundance, there was reasonable agreement between the two techniques for most miRNAs examined (Figure 3.2B). The main exceptions were c20399 (miR_t20) and c19166 (miR_ t124) which showed decreased abundance in the libraries from one or both photoautotrophic conditions relative to that from the mixotrophic condition whereas in the northern blots their steady-state levels remained relatively constant (Figures 3.2B). Additionally, the Maa7-IR44s and CC-124 strains generally showed comparable miRNA

levels, although unexpected differences were also observed for a few miRNAs (Figure 3.2, c16411 and c26753).

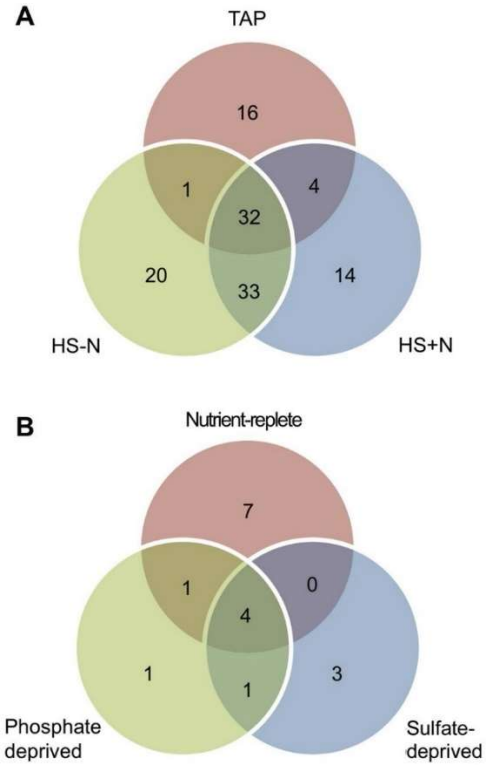

Figure 3.1. Comparison of miRNAs identified in Chlamydomonas cells grown under various nutritional deprivation conditions. (A) Venn diagrams show the numbers of unique and shared candidate miRNAs in cells grown under the different trophic regimes. The data was obtained from the AGO3-associated sRNA libraries in this study and (B) from the total sRNA libraries prepared by Chávez Montes *et al*.

To extend the pool of potentially functional miRNAs related to nutritional stress responses in *C. reinhardtii*, we also re-analyzed the sRNA data published by Chávez Montes *et al*.[25]. However, the libraries in their study were generated from total cellular small RNAs (under nutrient-replete, phosphate-deprived, or sulfur-deprived conditions), rather than from AGO3-associated sRNAs, and only 17 sequences met the criteria[14, 29] to be classified as canonical miRNAs in our analyses (Figure 3.1B and Table S1B). This

limited dataset nonetheless suggests that the C. reinhardtii miRNA population also varies among cells grown under phosphate or sulfur starvation although, as discussed above, when considering read abundance virtually none of the identified miRNAs is truly condition-specific (Table S1B). The expression of a subset of miRNAs in cells grown under phosphate or sulfur deprivation was also examined by RNA blotting and hybridization, but there was relatively poor agreement between the northern blotting signals (Figure S3A) and the total library read counts (Figure S3B). The studied cells did experience the expected nutritional deficiency, as indicated by the upregulation of diagnostic genes such as PHO5, encoding a phosphate-repressible alkaline phosphatase, and SLT1 (SAC1-LIKE TRANSPORTER1), encoding a sodium/sulfate cotransporter (Figure S4). Thus, the poor correlation in miRNA abundance between northern blotting and library read counts might be due to the fact that different Chlamydomonas strains were used for the analyses and/or technical issues (see Discussion). Nevertheless, our observations, taken together, indicate that several miRNAs are differentially expressed in response to nutrient depletion in Chlamydomonas, although very few (if any) appear to be strictly condition-specific.

Predicted miRNA targets in C. reinhardtii under various trophic conditions.

As previously described[14], potential miRNA targets were predicted based on sequence complementarity between miRNAs and binding sites on transcripts. In addition, predicted targets were classified, depending on the extent of complementarity to a miRNA, as likely to be regulated via transcript cleavage or translation repression[14]. As expected, changes in miRNA populations associated with each growth condition resulted in the prediction of distinct target gene populations (Figure 3.3).

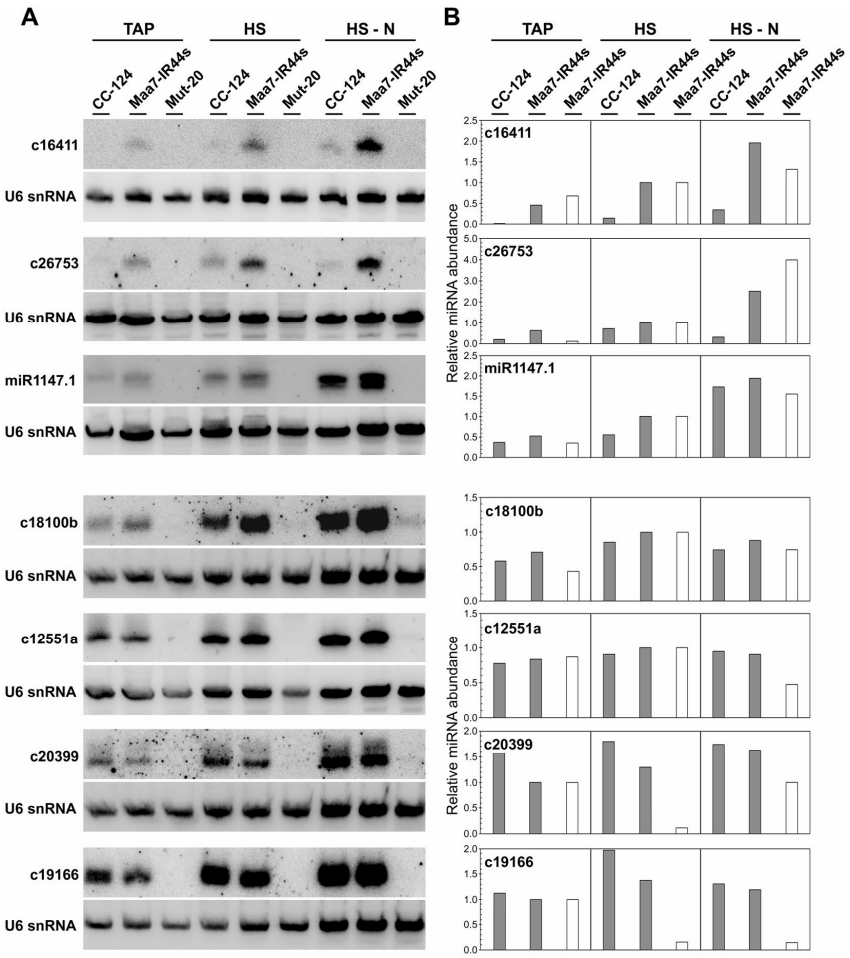

Figure 3.2. Northern blot analysis of miRNA expression in Chlamydomonas cells grown under the denoted trophic conditions. (A) Small RNAs were detected with probes specific for the indicated miRNAs. The same filters were reprobed with the U6 small nuclear RNA sequence as a control for lane loading. CC-124 wild type strain; Maa7-IR44s, CC-124 containing a transgene expressing FLAG-tagged AGO3; Mut-20, TSN1 deletion mutant, in the Maa7-IR44s background, defective in sRNA biogenesis14. (B) Relative miRNA levels in the indicated strains under the different trophic conditions. Values shown are the average of two independent experiments and are normalized to those of the Maa7-IR44s strain grown photoautotrophically in nutrient replete minimal medium (HS). For c20399 and c19166, values are normalized to those of Maa7-IR44s grown mixotrophically in acetate containing medium (TAP). The relative standard deviation, as percentage of the mean, was in no case higher than 28.3%. Data corresponds to phosphorimager measurements of sRNA signals on northern blots (gray bars) or normalized read counts from the AGO3-associated sRNA libraries (white bars).

To begin assessing whether miRNAs may play a regulatory role in the response to nitrogen starvation, we examined in more detail the predicted targets of the 20 miRNAs

that were more prevalent under nitrogen-deprived photoautotrophic conditions (Figure 3.1A, HS-N), as well as those of the 14 miRNAs that were identified primarily in nutrient replete photoautotrophic conditions (Figure 3.1A, HS + N). The miRNAs characteristic of HS − N had 58 putative cleavage targets and 207 translation repression targets whereas the miRNAs typical of HS + N potentially regulated 95 cleavage targets and 369 translational repression targets (Figure 3.3A and B, and Table S1A). However, the vast majority of the predicted targets corresponds to genes with unknown function and very few of those with an annotated function(s) code for proteins presumably involved in direct responses to nutrient deficiency (Table S2). Instead they appear to have a wide variety of cellular roles, including flagella associated proteins, molecular chaperones, protein kinases, post-translational modification proteins, predicted extracellular polypeptides and a few transcription factors (Table S2).

Transcriptome profiling revealed that of the 58 predicted cleavage targets for miRNAs more abundant in HS − N grown cells, only two showed at least a 2-fold decrease in steady-state mRNA levels under nitrogen starvation and a concomitant up-regulation in the miRNA-deficient Mut-20, as expected for true cleavage targets (Table S2, *Cre18.g749747* and *Cre06.g303200* highlighted in yellow). Likewise, of the 95 predicted cleavage targets for the 14 HS + N prevalent miRNAs, two displayed at least a 2-fold increase in expression in Mut-20 (Table S2, *Cre12.g552950* and *Cre16.g674291* highlighted in yellow). However, only one of these potential targets, *Cre12. g552950*, was differentially expressed during nitrogen starvation. Additionally, of the 207 potential translational repression targets for miRNAs more abundant in HS-N, only three showed a >2-fold change in transcript levels under nitrogen starvation as well as in the miRNA

deficient Mut-20 strain (Table S2, *Cre14.g627411*, *Cre07. g338000* and *Cre10.g464100* highlighted in yellow). Similarly, of the 369 predicted translational repression targets for the 14 miRNAs identified in HS + N, only 8 showed at least a 2-fold increase in transcript abundance in the miRNA deficient Mut-20 (Table S2, targets highlighted in yellow). Nonetheless, because miRNA regulation by translation inhibition does not necessarily alter the steady-state level of target transcripts[14, 16, 17], further analyses of protein abundance would be necessary to verify potential translation repression targets.

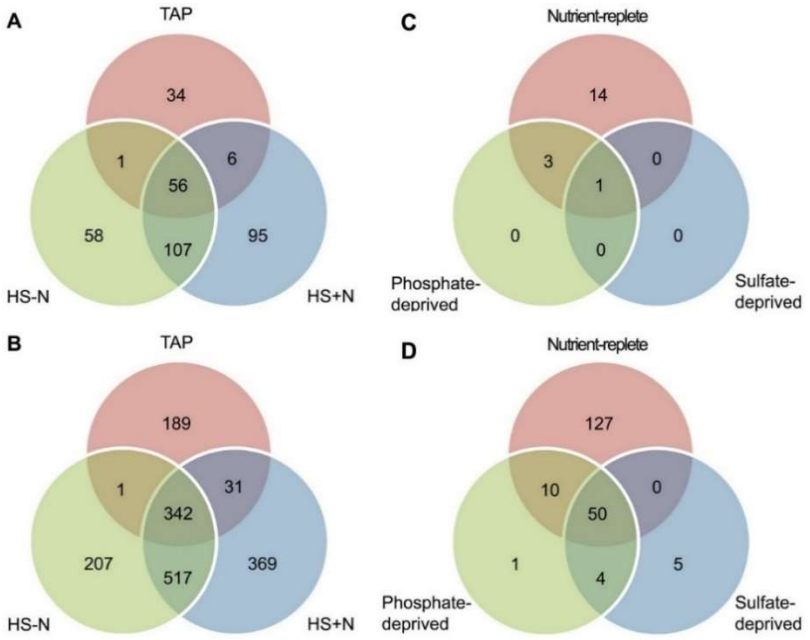

Figure 3.3. Comparison of predicted miRNA targets in Chlamydomonas cells grown under various nutritional deprivation conditions. Venn diagrams show the numbers of unique and shared putative miRNA targets in cells grown under the different trophic regimes. (A) and (C), Predicted cleavage targets. (B) and (D), Predicted translation repression targets. MicroRNA targets were computationally predicted (see Methods) based on the miRNAs identified from the AGO3-associated sRNA libraries in this study (A and B) and from the total sRNA libraries prepared by Chávez Montes *et al*. (**C** and **D**).

The transcript abundance of three predicted miRNA cleavage targets, up-regulated in the RNAseq experiments with Mut-20, was also verified by qRT-PCR analyses in Mut-20 and in a strain defective in AGO3, ago3-128 (see below), in comparison with their parental strains (Figure S5). We examined the putative targets of two miRNAs expressed at high levels (*Cre04.g227600* target of c12364 and *Cre06.g249550* target of c18100a) as well as the predicted target of one miRNAs expressed at low levels (*Cre12.g552950* target of miR_t70). *Cre12.g552950* behaved as a genuine cleavage target, with increased transcript abundance in both RNAi defective strains (Figure S5A). However, the *Cre12.g552950* mRNA is perfectly complementary to several sRNAs present in the libraries (including putative endogenous small interfering RNAs) (Figure S5B) and its steady state level may be modulated by the combined action of multiple sRNAs rather than solely by the lowly expressed miR_t70. *Cre04.g227600* and *Cre06.g249550* showed increased transcript levels only in the Mut-20 strain (Figure S5A). Interestingly, the c12364 and c18100a miRNAs are moderately reduced in abundance in ago3-1 (30–40% of wild type levels) and they may still be able to suppress *Cre04.g227600* and *Cre06.g249550* expression in this strain, in conjunction with Chlamydomonas AGO1 or AGO228. Of note, transcript abundance of all three target genes appears to be only modestly affected by the action of miRNAs/sRNAs (Figure S5A).

For the five miRNAs that were recognized preferentially during phosphate- and/or sulfur-deprivation (Figure 3.1B), we identified no potential cleavage targets and only 10 potential translational repression targets (Figure 3.3C and D, Table S1B), which were not differentially expressed under the conditions examined. Conversely, for the 7 miRNAs that were preferentially identified in total sRNA libraries from the nutrient-replete

condition (Figure 3.1B), we predicted 14 potential cleavage targets and 127 potential translational repression targets (Figure 3.3C and D, Table S1B). Of these putative targets, only one predicted cleavage target and three predicted translational repression targets were differentially expressed during phosphate- and/or sulfur-deprivation (Table S3). However, none of these genes codes for a protein involved in nutrient assimilation/metabolism and their putative role(s) in response to phosphate- and/or

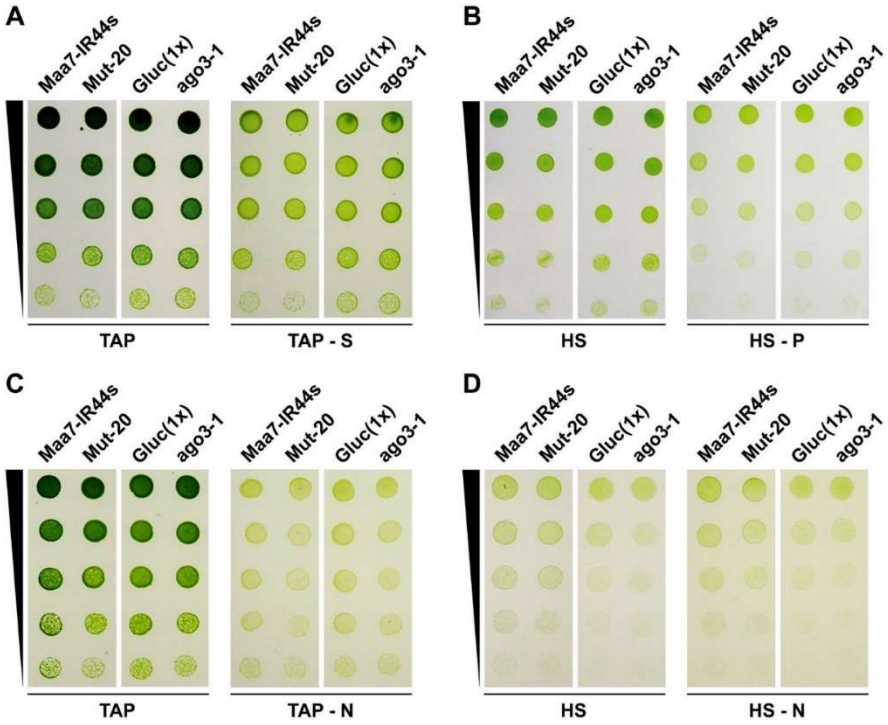

Figure 3.4. Growth and survival of Chlamydomonas cells subjected to various nutritional deprivation conditions. Cells grown to logarithmic phase in TAP medium were serially diluted in water, 5 µl-aliquots spotted on plates of the appropriate media and incubated for 7 to 15 days under continuous illumination. Maa7- IR44s, CC-124 strain containing a transgene expressing FLAG-tagged AGO3; Mut-20, TSN1 deletion mutant, in the Maa7-IR44s background, defective in sRNA biogenesis14; Gluc(1x), wild type strain derived from CC-124; ago3-1, AGO3 disrupted mutant, in the Gluc(1x) background, defective in RNA interference28. (A) Cells grown under mixotrophic conditions in the presence or absence of sulfur. (B) Cells grown under photoautotrophic conditions in the presence or absence of phosphorus. (C) Cells grown under mixotrophic conditions in the presence or absence of nitrogen. (D) Cells grown under photoautotrophic conditions in the presence or absence of nitrogen.

sulfur-deficiency is not clear. Moreover, in most cases, the changes in transcript abundance observed under nutrient deprivation (Table S3) were not in the expected direction based on the changes in abundance of the corresponding targeting miRNAs (Table S1B).

In summary, it remains uncertain how many of the predicted miRNA targets (under any of the examined trophic conditions) are genuine. However, even if some (many) predictions correspond to false positives, our observations strongly suggest that most targets in C. reinhardtii do not appear to be subject to miRNA-triggered transcript degradation, as reflected by the lack of changes in mRNA abundance in response to changes in miRNA abundance (in an sRNA-defective mutant strain or in cells exposed to various nutritional regimes inducing differential miRNA expression). As discussed below, we hypothesize that many Chlamydomonas miRNAs may be expressed at too low levels, under any trophic condition, to be functionally effective and those expressed at higher levels appear to have few, if any, targets.

Cell growth and survival of RNAi-defective C. reinhardtii strains under nutrient deprived conditions.

To examine further whether miRNAs may play a role in responses to nutrient depletion (and by inference in endogenous gene regulation), we tested the growth and survival of several RNAi defective strains under different trophic conditions. We assayed Mut-20, which contains a deletion of the gene coding for TUDOR STAPHYLOCOCCAL NUCLEASE1 (TSN1), implicated in sRNA biogenesis[14], and its parental strain Maa7-IR44s. We also analyzed ago3-1, which contains a disrupted AGO3 gene, and its parental

strain Gluc(1x)28. Mut-20 is virtually devoid of small RNAs14 whereas ago3-1 has reduced levels of a subset of miRNAs and defects in sRNA mediated post-transcriptional gene silencing[28]. Nonetheless, the growth of the mutant strains, under a variety of nutrient depletion conditions, was very similar to that of the parental strains (Figure 3.4).

We also examined cell survival after subjecting the strains to prolonged nitrogen-, phosphate- or sulfur-deprivation. Yet, for the most part, the mutants behaved like the wild type strains (Figure 3.5). The only sig- nificant difference was a moderate decrease in the survival of Mut-20, relative to Maa7-IR44s, upon exposure to nitrogen-depleted medium under photoautotrophic conditions (Figure 3.5A). However, it is not certain whether this reduced survival is due to a defect in miRNA-mediated gene regulation or caused by a deficiency in other pleotropic functions of the deleted TSN1 protein14, particularly since ago3-1 survival was not meaningfully compromised by nitrogen deprivation (Figure 3.5A). Thus, the lack of major phenotypic defects associated with disruption of the RNAi machinery (see also Valli *et al*.[30]) suggests a rather limited, modulatory role of miRNA-mediated gene regulation in Chlamydomonas cells cultured under nutrient deprived conditions.

Features of AGO3-associated *C. reinhardtii* miRNAs expressed under various trophic conditions. Many of the identified miRNAs have relatively low levels of expression (on average <500 Counts Per Million mapped reads or CPM) (Figure 3.6A, Table S1A). These miRNAs also tend to have a larger number of predicted targets than those with higher expression (Figure S6). In the most extreme case, a lowly expressed miRNA (miR_t69) was predicted to have 48 cleavage targets and 242 translational repression targets (Table S1A). When comparing low expression miRNAs (average

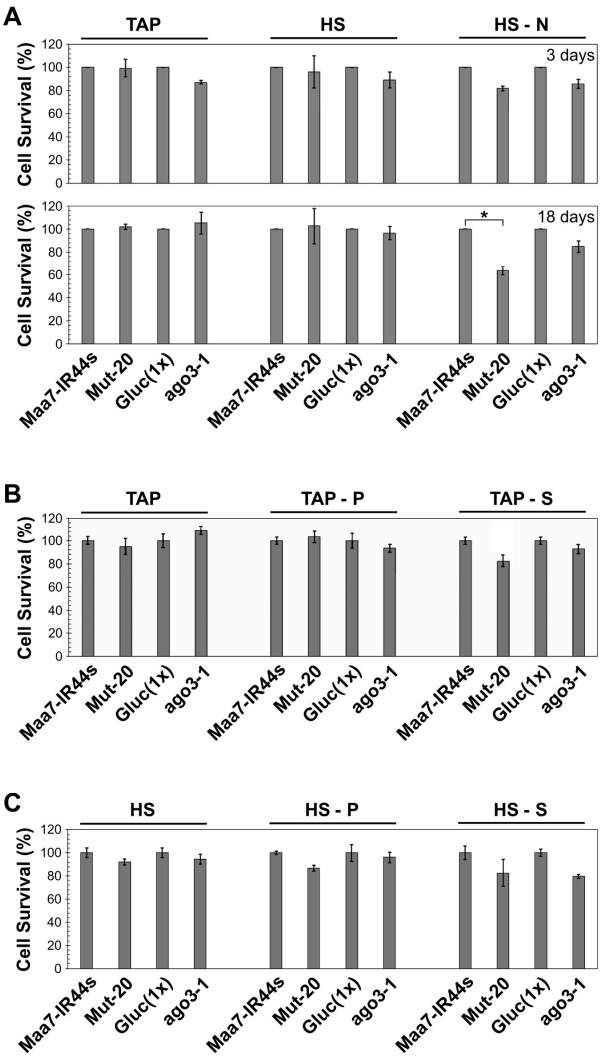

Figure 3.5. Viability of wild-type and RNA interference defective strains subjected to nutrient deprivation conditions. Cells were cultured in liquid medium, either replete or lacking a specific nutrient, for certain number of days and then spread on TAP-agar plates to assess survival as colony forming units. Values shown are the average of three independent experiments ± SD and are normalized to those of the control strains under each trophic condition. Maa7-IR44s, CC-124 strain containing a transgene expressing FLAG-tagged AGO3; Mut-20, TSN1 deletion mutant, in the Maa7-IR44s background, defective in sRNA biogenesis14; Gluc(1x), wild type strain derived from CC-124; ago3-1, AGO3 disrupted mutant, in the Gluc(1x) background, defective in RNA interference28. (A) Cell survival of the indicated strains grown mixotrophically (TAP) or photoautotrophically in the presence (HS) or absence (HS − N) of nitrogen for 3 or 18 days. Samples marked with an asterisk are significantly different ($p < 0.05$) in a two tailed Student's t-test. (B) Cell survival of the indicated strains grown mixotrophically for 18 days in nutrient replete medium (TAP) or lacking phosphorus (TAP-P) or sulfur (TAP-S). (C) Cell survival of the indicated strains grown photoautotrophically for 18 days in nutrient replete medium (HS) or lacking phosphorus (HS-P) or sulfur (HS-S).

expression <500 CPM) classified as miRNAs only in cells grown under a certain nutritional regime (TAP, HS + N or HS − N) with high expression miRNAs (average expression >500 CPM) shared under all nutritional conditions, the average number of predicted targets per miRNA was significantly different (Figure 3.7). The lowly expressed miRNAs have an average of 5.11 (n = 47) predicted cleavage targets per miRNA, whereas highly expressed miRNAs have an average of 1.78 (n = 18) predicted cleavage targets per miRNA (p = 0.0156 by Wilcoxon rank sum test, Cohen's d = 0.538) (Figure 3.7). This trend was also observed for predicted translational repression targets, but with smaller (non-significant) differences. Because non-conserved, recently evolved miRNAs in higher plants tend to have low expression levels[24–26, 31, 32], similarly to many Chlamydomonas miRNAs, we performed equivalent analyses on miRNAs specific to either *A. thaliana* or *A. lyrata*. Most species-specific miRNAs in each Arabidopsis species are lowly expressed (<500 CPM) (Figure 3.6B and C). They also tend to show an inverse relationship between miRNA expression level and number of predicted targets (Figures 3.7 and 3.8A). The difference was more prominent in *A. thaliana*, with an average of 9.25 (n = 40) predicted cleavage targets per lowly expressed miRNA and only 2.75 (n = 4) predicted cleavage targets per highly expressed miRNA (Figure 3.7). However, likely due to the small number of highly expressed miRNAs, this difference was not statistically significant (p = 0.075 by Wilcoxon rank sum test). A similar trend was observed in A. lyrata, with an average of 4.73 (n = 34) predicted cleavage targets per lowly expressed miRNA and 2.80 (n = 5) predicted cleavage targets per highly expressed miRNA (Figure 3.7) although, as in *A. thaliana*, the difference was not statistically significant (p = 0.915 by Wilcoxon rank sum test). In contrast, miRNAs that are

conserved between *A. thaliana* and *A. lyrata* (many with experimentally validated roles) are expressed at higher levels[26, 32] and have on average a greater number of predicted cleavage targets per miRNA (Figures 3.7 and 3.8B, C). Moreover, by target degradome sequencing, nearly all verified cleavage targets in A. thaliana and A. lyrata were those corresponding to conserved miRNAs[26].

We also examined conservation of the Chlamydomonas miRNAs (Table S1) against all miRNAs (either mature or stem-loop sequences) deposited in miRBase[33]. The search revealed no significant hits to miRNAs from any organism aside from *C. reinhardtii* (see Methods). Previous studies also showed that the miRNAs identified in Chlamydomonas had no meaningful similarity to mature miRNA sequences even from the alga *Volvox carteri*[14, 18, 19, 34], the closest relative to C. reinhardtii for which sRNAs and miRNAs have been extensively profiled. Thus, Chlamydomonas miRNAs are not evolutionarily conserved and many show relatively low expression levels (particularly those identified primarily under certain nutritional conditions) as well as negative correlation between miRNA abundance and number of predicted targets. These features largely resemble those of the species-specific, recently evolved miRNAs characterized in higher plants[24–26, 31, 32].

Discussion

Comparison of AGO3-associated sRNA populations from Chlamydomonas cells grown under various trophic conditions revealed that some miRNAs are differentially expressed in response to nutritional changes, but none of the identified miRNAs appears

to be strictly condition specific (Table S1A). The re-analysis of total sRNA libraries prepared by Chávez Montes *et al*.[25] from *C. reinhardtii* cultured under nutrient replete, phosphate-deprived or sulfur-deprived conditions also supported the existence of differentially expressed miRNAs as a result of nutrient depletion (Table S1B). Changes in steady-state levels for a subset of the miRNAs were corroborated by northern blot analyses (Figures 3.2 and S3), although there was better agreement, with some exceptions, between northern blot signals and normalized read counts from the AGO3-associated sRNA libraries than from the total sRNA libraries. This may be explained by experimental variation since different Chlamydomonas strains were used in the latter comparison (see Methods and Chávez Montes *et al*.[25]).

As previously demonstrated, the representation of miRNA sequences relative to each other within a sRNA library may not be consistent with their input concentrations, because of biases in ligation-based small RNA library construction due to adaptors, RNA structure, and RNA ligase activity[35–38]. On the other hand, this problem is expected to be less relevant (i.e., a systematic bias) for relative abundance comparisons of the same miRNA across libraries prepared in the same way. Nonetheless, for some miRNAs this may still be a problem since their relative levels in sRNA libraries from various trophic conditions differed substantially from their detection by northern blotting. From a technical perspective, it seems clear that library construction (either from total or AGO-associated sRNAs) does affect the identification of potentially functional miRNAs. However, our combined observations, based on the analyses of multiple sRNA libraries as well as northern blotting, strongly support that changing nutritional conditions induces the differential expression of a subset of miRNAs in *C. reinhardtii*.

Since AGO3 is the main effector of sRNA-mediated post-transcriptional gene silencing in Chlamydomonas[28], changes in AGO3-associated miRNAs may potentially be of functional relevance in responses to nutritional stress. Thus, we surveyed the putative role(s) of computationally predicted targets for the 20 miRNAs that were more prevalent under nitrogen-deprived photoautotrophic conditions (Figure 3.1A, HS - N) and for the 14 miRNAs that were identified preferentially in nutrient replete photoautotrophic conditions (Figure 3.1A, HS + N). However, most predicted targets corresponded to genes with unknown function, and virtually none of those with an annotated function(s) encoded a protein directly involved in nitrogen metabolism/assimilation (Table S2). Additionally, transcriptome profiling of cells cultured in nutrient-replete or in nitrogen-deprived media as well as of a mutant strain, Mut-20, virtually devoid of sRNAs revealed that very few of the putative miRNA targets showed changes in transcript abundance consistent with their regulation by miRNA-mediated RNA degradation (Table S2). Similar observations were made for cells grown under phosphate- or sulfur-deprived conditions (Table S3).

The Chlamydomonas RNAi machinery has the capability to operate by target transcript cleavage, as demonstrated with artificial miRNA transgenes[16, 28, 39, 40]. However, with the caveat that some predictions may represent false positives, most endogenous miRNA targets do not appear to be subject to transcript cleavage and degradation in cells cultured under multiple trophic conditions. A similar conclusion was reached by Valli *et al*.[30] through the analysis of a Chlamydomonas mutant defective in DICER LIKE3 (DCL3), which failed to produce both miRNAs and siRNAs. Nonetheless, the Chlamydomonas RNAi machinery also has the capability to cause translation

repression of target transcripts, as demonstrated with transgenic constructs[16, 17, 28]. MicroRNAs could also exert regulatory roles on host transcripts in cis, simply by being processed by Dicer, since several Chlamydomonas miRNAs are derived from mRNAs of hypothetical protein coding genes[14, 30].

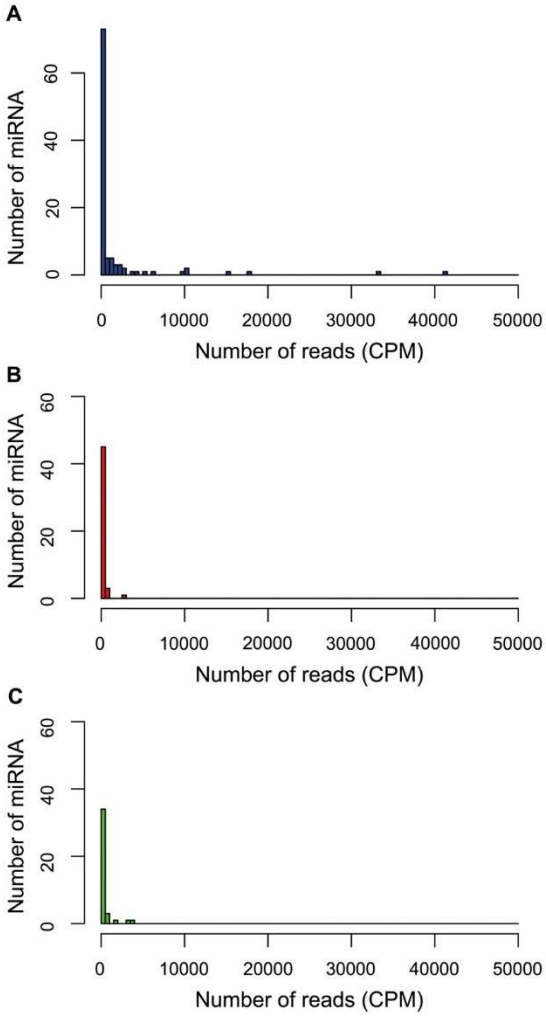

Figure 3.6. Distribution of miRNAs based on their average expression levels for C. reinhardtii. (A) and for the species-specific miRNAs in Arabidopsis thaliana (B) and Arabidopsis lyrata (C). Binning was done with 500CPM (Counts Per Million mapped reads) intervals.

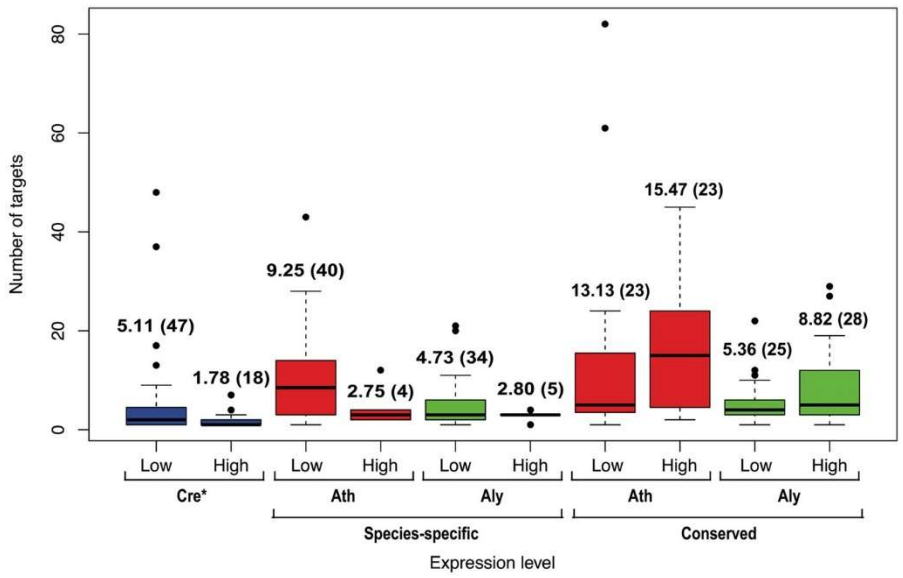

Figure 3.7. Comparison of the number of predicted cleavage targets for lowly expressed and highly expressed miRNAs. Boxplots show the number of predicted cleavage targets for miRNAs with low expression (<500 CPM) and high expression (>500 CPM) identified in the AGO3-pulldown libraries from C. reinhardtii (blue) and for species-specific or conserved miRNAs from Arabidopsis thaliana (red) or Arabidopsis lyrata (green). Conserved miRNAs are shared between A. thaliana and A. lyrata and many have experimentally validated roles. MicroRNAs that had no predicted targets were excluded from this plot. The average number of targets (including outliers) and the sample size for each category are shown above the boxes. The target numbers were compared between the two expression-level groups for each species and an asterisk next to the species abbreviation indicates a significant difference ($p < 0.05$ by Wilcoxon rank sum test).

Conceivably, miRNAs differentially expressed under various trophic conditions could play an important role in responses to nutritional stress, via modulation of target translation efficiency, the stability of some host transcripts or even some unconventional mechanisms[7, 41–43]. However, this interpretation is not supported by phenotypic analyses of strains defective in components of the RNAi machinery. Mut-20 is virtually devoid of sRNAs[14] whereas ago3-1 has reduced levels of a subset of miRNAs and defects in sRNA mediated post-transcriptional gene silencing[28]. Despite these major alterations to the RNAi machinery both mutants grew as well as the parental strains under mixotrophic or

photoautotrophic conditions, in liquid or solid medium, and under various nutrient deprived conditions (Figure 3.4). As already mentioned, even the moderate decrease in Mut-20 survival after pro- longed exposure to nitrogen depleted minimal medium (Figure 3.5A) cannot be unequivocally ascribed to a defect in miRNA-mediated gene regulation. Likewise, the Chlamydomonas dcl3 mutant did not show obvious alterations in growth or morphological abnormalities under normal laboratory conditions[30, 44]. Moreover, ribosome profiling and proteomic analyses in the wild type and the dcl3 mutant indicated that miRNAs have little effect on translation efficiency and largely fine tune target gene expression[44]. Thus, while a recent report proposed that certain miRNAs play a key role in abiotic stress responses in *C. reinhardtii*[10], accumulating evidence suggests that most miRNAs mainly have a modulatory, rather modest function in the regulation of biological processes in this alga (at least under normal and nutrient deprived growth conditions). This hypothesis is very difficult to demonstrate conclusively since it can only be supported by negative data (i.e., the lack of a verified miRNA function) but it seems the most parsimonious explanation for the collective results of us and others[14, 28, 30, 44].

In addition, Chlamydomonas miRNAs are not evolutionarily conserved even within the order Volvocales, which includes the related alga Volvox carteri[14, 18, 19, 34]. Chlamydomonas and Volvox lineages diverged ~200 mil- lion years ago[45] and extensive sequence divergence over this length of evolutionary time may have obscured miRNA homologies. However, this seems unlikely to be the case for all miRNA loci, since subsets of both animal and land plant miRNAs have been strongly conserved over a similar period of time[20, 25, 29, 32, 46, 47]. Thus, Chlamydomonas miRNAs appear to have evolved relatively recently, since the divergence from the lineage leading to the family

Volvocaceae. Moreover, many Chlamydomonas miRNAs, particularly those identified pref- erentially under certain nutritional conditions, are expressed at relatively low levels (Table S1A and Figure 3.6A) and show negative correlation between miRNA abundance and number of predicted targets (Figures 3.7 and S6). Many condition-prevalent miRNAs also seem to show imprecise processing from fairly long hairpin precursors (Figure S1), which is reflected in lower predominance of the reads corresponding exactly to the mature miRNAs, representing <90% of all the reads mapping locally to the precursor hairpins. These features resemble those of the species-specific, newly evolved miRNAs characterized in land plants[24–26, 31, 32, 47, 48].

Recent findings in metazoans suggest that only strongly expressed miRNAs, above a certain threshold level, may lead to functionally significant target suppression. By using a sensor library to monitor miRNA activity in human monocytes, only miRNAs expressed above 100-1000 reads per million showed suppressive activity[49]. High miRNA abundance might be necessary to facilitate miRNA interaction with target transcripts, through diffusion and sampling within a cell, although the extent of suppression also depends, among other variables, on target site concentration[49–51]. Assuming a similar expression threshold for functional miRNAs in Chlamydomonas, we hypothesize that over 60% of the AGO3-associated miRNAs (detected at <500 CPM under any condition, Table S1A) would not be expected to have discernable activity. This is consistent with expectations for young miRNAs since an initial weak expression and negligible fitness effects would allow their progressive integration into gene regulatory networks[32, 48, 52].

The RNAi machinery presumably arose as an ancestral defense mechanism against selfish genetic elements such as viruses and transposons[53–55] and was later

coopted to miRNA pathways that evolved independently in several eukaryotic lineages[2, 3, 47, 54, 55]. As proposed in a number of organisms[2, 20, 32, 47, 52], low level transcription of inverted repeats or mutationally engendered hairpin structures could give rise to a diversity of RNAs recognized as substrates by the sRNA biogenesis machinery.

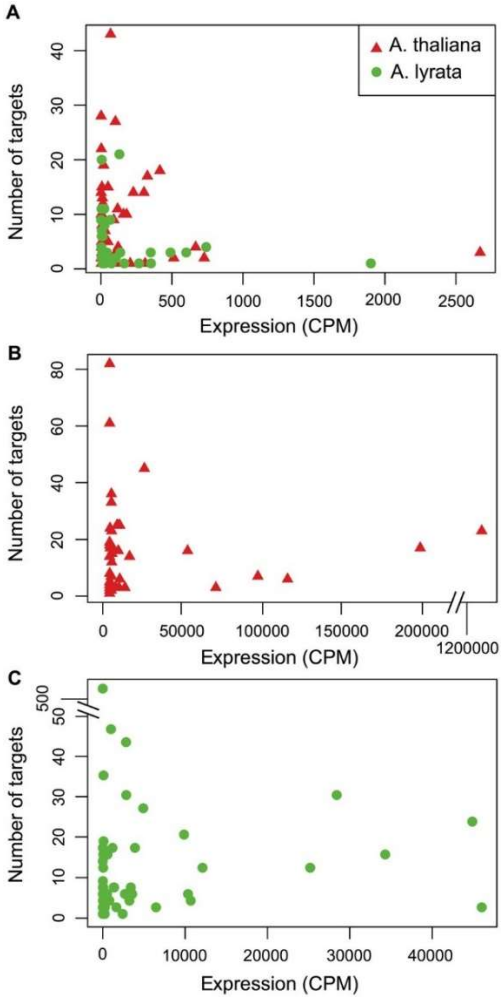

Figure 3.8. Relationship between miRNA expression level and number of predicted targets in Arabidopsis thaliana and Arabidopsis lyrata. Scatter plots show the expression level and the number of predicted targets for each miRNA in A. thaliana (red) and A. lyrata (green). (A) Species-specific miRNAs in both Arabidopsis species. (B) Conserved miRNAs in A. thaliana. (C) Conserved miRNAs in A. lyrata.

However, most young miRNAs would likely be neutral[32, 47, 48, 52], either by not being expressed at a high enough level or by not having enough sequence identity to regulate any meaningful target. Random mutations and genetic drift would lead to the relatively rapid evolutionary turnover of these miRNA precursor genes. In contrast, miRNAs that acquire a target with functional relevance would be maintained under purifying selection and could increase, over time, their expression and even acquire additional targets to enable more efficient gene regulation[52]. Accordingly, in both land plants and animals, conserved, older miRNAs are generally expressed at higher levels and have more targets than young ones[23–26, 32, 47, 48, 52].

The eukaryotic groups that exhibit the highest level of multicellular complexity (animals and land plants) all possess miRNAs[2, 20, 25, 29, 46, 47, 55] that regulate important biological processes, including cell differentiation and development[1, 3, 7, 32, 46, 47]. This correlation has led several authors to propose that miRNAs may have played a role in the evolution of complex multicellularity[55–57]. In contrast, in the unicellular alga *C. reinhardtii* the miRNA system appears to consist largely of recently evolved miRNAs that, based on the RNAi-defective mutant phenotypes, do not seem to play a substantial role in cell growth and survival (at least under the trophic conditions examined). Indeed, Chlamydomonas miRNAs do not appear to have been meaningfully integrated yet into the organism's gene regulatory network. Even in eukaryotes at an early transition towards a multicellular stage, such as *Volvox carteri* and *Dictyostelium discoideum*, the role of miRNAs in controlling gene expression is elusive[19, 58]. For instance, *D. discoideum* drnB-mutant cells, lacking a Dicer-like protein required for miRNA biogenesis, grow and develop normally[58]. Thus, it is tempting to speculate that in unicellular eukaryotes,

miRNAs arising as accidental products of random genome evolution may provide no major selective advantage for the regulation of essential cellular functions, ancestrally controlled by other components such as transcription factors. Most miRNAs in these organisms may be transient, without a main biological utility, but may provide a pool from which new miRNA-target regulatory interactions could eventually be recruited leading to evolutionary innovations.

Methods

Strains, mutants, and culture conditions

Chlamydomonas cells were grown mixotrophically in TAP medium[17] or photoautotrophically in high salt (HS) medium[59]. For nitrogen deprivation analyses, cells initially grown photoautotrophically in nutrient replete medium to the middle of the logarithmic phase were collected by centrifugation and resuspended at a density of ~1.0 × 106 cells mL−1 in the same medium with (HS + N) or without nitrogen (HS − N). After 72 h of incubation under continuous illumination (180 μmol m−2 s−1 photo-synthetically active radiation), cells were harvested and immediately frozen in liquid nitrogen for subsequent RNA isolation or FLAG-tagged AGO3 purification. A similar protocol was used for the analysis of phosphate or sulfur deprived cells, following prior specifications[15, 25]. The wild type strain, CC-124, and a transgenic strain, Maa7-IR44s, containing an inverted repeat construct targeting the 3' UTR of the MAA7 gene (encoding tryptophan synthase β subunit) and the FLAG-tagged AGO3, have been previously described[14, 17]. Mut-20, deleted for the TSN1 gene, was obtained in an

insertional mutagenesis screen designed to isolate mutants defective in RNAi-mediated translation repression[14]. Likewise, ago3-1, containing a disrupted AGO3 gene, was isolated in a forward mutagenesis screen in the Gluc(1x) background[28]. For estimating strain survival under nutritional stress, cells were grown in TAP or HS medium to the middle of the logarithmic phase, washed three times in the desired medium, and resuspended to a density of ~1.0 × 106 cells mL−1 in nutrient replete TAP or HS medium or in the same medium lacking N, P or S. After incubation in liquid medium, under standard culture conditions, for 3 or 18 d, aliquots of cells were spread on TAP-agar plates (5 replicates per treatment and strain) to assess survival as colony forming units.

Isolation of AGO3-associated sRNAs, library preparation, and sequencing

FLAG-tagged AGO3 was affinity purified from cell lysates as previously described for a TAP-tagged protein60. RNAs associated with AGO3 were purified with TRI reagent (Molecular Research Center) and contaminant DNA was removed by DNase I treatment (Ambion)[17]. Construction of cDNA libraries and Illumina sequencing were then carried out as previously reported[61]. AGO3-associated sRNAs were characterized from cells grown mixotrophically in TAP or photoautotrophically in HS + N or HS − N media (NCBI accession numbers SRR1747077, SRR2959984 and SRR2959993, respectively).

sRNA mapping and profiling

Sequenced reads were first mapped to the C. reinhardtii genome[62], by using version 3.02 of Novoalign (www.novocraft.com) with the miRNA flag and with a score

threshold of 15. Mapped reads were filtered to remove those showing alignments with gaps or mismatches as well as those that mapped to more than five locations in the genome. Reads mapping to the chloroplast or mitochondrial genomes or to functional non-coding RNAs were also removed, as previously described[14]. The expression level in counts per million (CPM) for each mapped sRNA was determined by the formula:

$$CPM = [(106C)/N]$$

where C is the number of mapped reads corresponding to an individual sRNA sequence in the library and N is the total number of mapped reads in the library. We also re-analyzed, in the same manner, the sRNA libraries generated by Chávez Montes *et al*.[25] (accession number GSM803103) to identify miRNAs related to phosphate or sulfur deprivation.

Genomic clustering of sRNAs and miRNA identification

Clusters of reads were identified as previously described[14] and the genomic sequence for each strand of a cluster was folded using version 2.1.5 of RNAfold from the Vienna RNA package[63]. Clusters containing sequence gaps (i.e., unsequenced genomic regions) were excluded from further analyses since the secondary structure of these regions cannot be unambiguously predicted. The obtained secondary structures were then parsed to determine if they fold into a hairpin. Clusters remaining after this filtering were manually curated based on the processing accuracy of the 5' end of the predominant read(s), the frequency of the predominant read(s), and the extent of complementarity

between the two arms of the hairpins, according to the criteria for canonical miRNA prediction[14, 29].

RNA analyses

Total RNA was isolated with TRI reagent (Molecular Research Center, Inc.)[17, 61], in accordance with the manufacturer's instructions, from *C. reinhardtii* cells grown under the different trophic conditions. The same RNA samples were used for northern blotting and for transcriptome analyses (see below). For sRNA northern analyses, total RNA samples were resolved in 15% polyacrylamide/7-M urea gels and electroblotted to Hybond-XL membranes (GE Healthcare)61. Blots were hybridized with 32P-labeled DNA probes using the High Efficiency Hybridization System at 40 ºC for 72 h17, 61. Specific miRNAs were detected by hybridization with DNA oligonucleotides labeled at their 5' termini with [α-32P]ATP and T4 Polynucleotide Kinase (New England Biolabs)[17, 61]. For quantitative RT-PCR analyses, DNase I-treated RNA samples were used as template for first-strand cDNA synthesis, using an oligo(dT)18 primer and SuperScript III reverse transcriptase (Life Technologies). Primer pairs for the quantitative PCR amplifications were as follows: for *Cre04.g227600*, LRR-F (5'-ACCCATGCTCT AAGGACTGGA-3') and LRR-R (5' - GTCGGAGAAGCAGGTGA GTGT-3'); for *Cre06.g249550*, 249550-F (5' -GGGAAAGAGTGGATGATGTGG-3') and 249550-R (5'-ACATCAACGTTGTGCCTCACT-3'); and for *Cre12.g552950*, 552950-F (5'-AACTGGATAGGCTGAGCAGGA-3') an d 552950-R (5'-TTGTGGGGACAGCT TCTTCTT-3'). The ACTIN1 transcript17, 61 was amplified for normalization purposes. DNA fragments were amplified and quantified with the RT2 SYBR Green/Fluorescein

qPCR mastermix (Qiagen), using the iCycler Real Time PCR Detection System (Bio-Rad). For semi-quantitative RT-PCR, the number of cycles showing a linear relationship between input cDNA and the final product were deter- mined in preliminary experiments[17]. Aliquots of each RT-PCR were resolved on 1.2% agarose gels and visualized by ethidium bromide staining. The primer sequences were as follows: for PHO5 (*Cre04.g216700*), PHO5-5F (5'-TTCCGTTTCCGTTCTCTGAC-3') and PHO5-3R (5'-CCCTGCATCTTGTTCTCCAG-3'); for SLT1 (*Cre12.g502600*), SLT1-5F (5'-ACGGGT TCT TCGAGCGAAT TGC-3') and SLT1-3R (5'-CGACTGCT TACGCA ACAATCTTGG-3'); for CBLP (*Cre06.g278222*), CBLP-5F (5'-CTTCTCGCCCATG ACCAC-3') and CBLP-3R (5'-CCCACCAGGTTGTTCTTCAG-3'); and for the U6 snRNA, U6-F (5'-TGCTTCGGCACAACTGTTAAA-3') and U6-R (5'-AAAATTTGG AACCATTTCTCGATT-3').

MicroRNA target prediction

Potential miRNA-binding sites in transcripts were determined as previously described[14], by searching v11 of the Phytozome *C. reinhardtii* transcriptome using version 2.1 of RNAhybrid[64]. For cleavage targets, this search required perfect matching for nucleotides 2–8 (the miRNA seed region) and nucleotides 9–12 (the miRNA catalytic center), and no more than three G:U wobbles and three mismatches or a gap of >1 nt in the remaining sequence. For translational repression targets, the constraints for the catalytic region were relaxed to allow up to three mismatches or wobbles. Additionally, translational-repression targets needed at least one mismatch or wobble in the catalytic region to keep the two sets of predicted targets non-overlapping[14]. Putative functions of

the predicted targets were evaluated by using the annotations of Chlamydomonas genes (if available) as well as conserved protein domains. Functional annotations were obtained with the Algal Functional Annotation Tool[65] and are mostly based on those in Phytozome v11[66].

Differential gene expression analyses

Transcriptome sequencing was performed on RNA samples isolated from Maa7-IR44s and Mut-20 grown photoautotrophically in HS + N or HS − N media or mixotrophically in TAP medium (NCBI accession numbers SRX1451698, SRX1451708, and SRR1747017, respectively). Experiments were performed twice, independently, and libraries were sequenced with the Illumina GAIIx analyzer, as previously described[14]. Illumina reads were mapped to the Augustus v5.0 transcript models for *C. reinhardtii* (available from http://genome.jgi-psf.org/Chlre4/Chlre4.download.ftp.html), by using Burrows-Wheeler Aligner (BWA; v0.5.7)[67] with a seed length of 25 and allowing 2 mismatches. An in-house Perl script was used to ensure that only reads that matched uniquely to a single transcript were counted. Raw gene counts were determined by adding the number of reads aligned to each transcript. RNA-Seq data for sulfur- or phosphate-deprived samples were taken from Gonzalez-Ballester *et al*.[15] and Schmollinger *et al*.68, respectively (accession numbers GSE17970 and GSE56505). Transcript abundance was analyzed as Reads Per Kilobase of transcript per Million mapped reads (RPKM), which normalizes read counts based on both transcript length and total number of reads, using the formula:

$$RPKM = [(109C)/(NL)]$$

where C is the number of reads mapped to each transcript, N is the total number of mapped reads in the library, and L is the transcript length in nucleotides[69]. To assess changes in gene expression, transcript abundance was compared between Mut-20 and its parental strain Maa7-IR44s, under each trophic condition, or between different nutritional conditions for the same strain. Differences in gene expression were examined as log2(FC), where FC (Fold Change) refers to the ratio of RPKM values between compared strains or treatments. Statistical analysis of the data was performed using the DESeq package (version 1.18)[70]. Genes with a q-value < 0.05 and >2-fold change in transcript abundance under at least one of the pairwise comparisons were considered differentially expressed. The Augustus v5.0 transcript IDs were converted to the Phytozome v11 transcript IDs using the name conversion file on the Phytozome website (https://phytozome.jgi.doe.gov/pz/portal. html#!bulk?org=Org_Creinhardtii).

Arabidopsis miRNAs and target prediction

Predicted miRNAs and their expression information for *A. thaliana* and *A. lyrata* were taken from Ma *et al*.[26]. Expression levels, given in raw read counts, were converted to CPM as described above. Species-specific miRNAs were determined by comparing the miRNA datasets for *A. thaliana* and *A. lyrata* and cross-referencing with in-text results[26]. Targets for the miRNAs were predicted using version 1.6 of Target Finder[31], searching against the *A. thaliana* and the *A. lyrata* transcriptomes taken from Phytozome v11[66].

Comparisons of miRNA expression and number of predicted targets. All statistical analyses were performed in R using standard libraries. The histograms of miRNA expression were generated by binning miRNAs, based on their average CPM levels, at 500 CPM intervals. MicroRNAs that had no identifiable target or only targeted their precursor transcript were excluded from further analyses, since these miRNAs would not have constraints on their expression level. A cutoff of 500 CPM, chosen based on the findings of Mullokandov *et al*.[49] for functionally effective miRNAs in metazoans, was used to classify lowly expressed (presumably non-functional) and highly expressed (potentially functional) miRNAs. The numbers of targets predicted for highly expressed and lowly expressed miRNAs were compared using Wilcoxon rank sum test, and Cohen's d was used to determine the effect size for the two groups.

Analysis of conservation of Chlamydomonas miRNAs

The identified Chlamydomonas mature miRNA sequences were compared against both the mature miRNA and the pre-miRNA hairpin sequences in release[21] of miRBase[33]. Both sequence similarity searches were performed using version 2.2.30+ of blastn with an e-value cutoff of 10^{71}. To increase the chances of finding conserved miRNAs, the search was performed against the entire database rather than limiting it to the high confidence miRNAs.

References

1. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function.

Nat Rev Mol Cell Biol 14, 475–488 (2013).

2. Cui, J., You, C. & Chen, X. The evolution of microRNAs in plants. Curr Opin Plant Biol 35, 61–67 (2016).

3. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. Cell 136, 215–233 (2009).

4. Zhao, T. *et al*. A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes Dev 21, 1190–1203 (2007).

5. Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C. & Baulcombe, D. C. miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447, 1126–1129 (2007).

6. Shu, L. & Hu, Z. Characterization and differential expression of microRNAs elicited by sulfur deprivation in Chlamydomonas reinhardtii. BMC Genomics 13, 108 (2012).

7. Wang, H. L. & Chekanova, J. A. Small RNAs: essential regulators of gene expression and defenses against environmental stresses in plants. Wiley Interdiscip Rev RNA 7, 356–381 (2016).

8. Nguyen, G. N., Rothstein, S. J., Spangenberg, G. & Kant, S. Role of microRNAs involved in plant response to nitrogen and phosphorous limiting conditions. Front Plant Sci 6, 629 (2015).

9. Khraiwesh, B., Zhu, J. K. & Zhu, J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta
1819, 137–148 (2012).

10. Gao, X. *et al*. MicroRNAs modulate adaption to multiple abiotic stresses in Chlamydomonas reinhardtii. Sci Rep 6, 38228 (2016).

11. Liang, G., He, H. & Yu, D. Identification of nitrogen starvation-responsive microRNAs in Arabidopsis thaliana. PLoS One 7, e48951 (2012).

12. Shukla, L. I., Chinnusamy, V. & Sunkar, R. The role of microRNAs and other endogenous small RNAs in plant stress responses.
Biochim Biophys Acta 1779, 743–748 (2008).

13. Ding, Y. F. & Zhu, C. The role of microRNAs in copper and cadmium homeostasis. Biochem Biophys Res Commun 386, 6–10 (2009).

14. Voshall, A., Kim, E. J., Ma, X., Moriyama, E. N. & Cerutti, H. Identification of AGO3-Associated miRNAs and Computational Prediction of Their Targets in the Green Alga Chlamydomonas reinhardtii. Genetics 200, 105–121 (2015).

15. Gonzalez-Ballester, D. *et al*. RNA-seq analysis of sulfur-deprived Chlamydomonas cells reveals aspects of acclimation critical for cell survival. Plant Cell 22, 2058–2084 (2010).

16. Yamasaki, T. *et al*. Complementarity to an miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga Chlamydomonas reinhardtii. Plant J 76, 1045–1056 (2013).

17. Ma, X. *et al*. Small interfering RNA-mediated translation repression alters ribosome sensitivity to inhibition by cycloheximide in Chlamydomonas reinhardtii. Plant Cell 25, 985–998 (2013).

18. Evers, M., Huttner, M., Dueck, A., Meister, G. & Engelmann, J. C. miRA: adaptable novel miRNA identification in plants using small RNA sequencing data. BMC Bioinformatics 16, 370 (2015).

19. Li, J., Wu, Y. & Qi, Y. MicroRNAs in a multicellular green alga Volvox carteri. Sci China Life Sci 57, 36–45 (2014).

20. Nozawa, M., Miura, S. & Nei, M. Origins and evolution of microRNA genes in plant species. Genome Biol Evol 4, 230–239 (2012).

21. Loh, Y. H., Yi, S. V. & Streelman, J. T. Evolution of microRNAs and the diversification of species. Genome Biol Evol 3, 55–65 (2011).

22. Barakat, A., Wall, P. K., Diloreto, S., Depamphilis, C. W. & Carlson, J. E. Conservation and divergence of microRNAs in Populus.
BMC Genomics 8, 481 (2007).

23. Felippes, F. F., Schneeberger, K., Dezulian, T., Huson, D. H. & Weigel, D. Evolution of Arabidopsis thaliana microRNAs from random sequences. RNA 14, 2455–2459 (2008).

24. Fahlgren, N. *et al*. MicroRNA gene evolution in Arabidopsis lyrata and Arabidopsis thaliana. Plant Cell 22, 1074–1089 (2010).

25. Chávez Montes, R. A. *et al*. Sample sequencing of vascular plants demonstrates widespread conservation and divergence of microRNAs. Nat Commun 5, 3722 (2014).

26. Ma, Z., Coruh, C. & Axtell, M. J. Arabidopsis lyrata small RNAs: transient MIRNA and small interfering RNA loci within the Arabidopsis genus. Plant Cell 22, 1090–1103 (2010).

27. Sharma, N., Tripathi, A. & Sanan-Mishra, N. Profiling the expression domains of a rice-specific microRNA under stress. Front Plant Sci 6, 333 (2015).

28. Yamasaki, T., Kim, E. J., Cerutti, H. & Ohama, T. Argonaute3 is a key player in miRNA-mediated target cleavage and translational repression in Chlamydomonas. Plant J 85, 258–268 (2016).

29. Tarver, J. E., Donoghue, P. C. & Peterson, K. J. Do miRNAs have a deep evolutionary history? BioEssays 34, 857–866 (2012).

30. Valli, A. A. *et al*. Most microRNAs in the single-cell alga Chlamydomonas reinhardtii are produced by Dicer-like 3-mediated cleavage of introns and untranslated regions of coding RNAs. Genome Res 26, 519–529 (2016).

31. Fahlgren, N. *et al*. High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes.
PLoS One 2, e219 (2007).

32. Cuperus, J. T., Fahlgren, N. & Carrington, J. C. Evolution and functional diversification of MIRNA genes. Plant Cell 23, 431–442 (2011).

33. Kozomara, A. & Griffiths-Jones, S. miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res
42, D68–D73 (2014). (Database issue).

34. Dueck, A. *et al*. Gene silencing pathways found in the green alga Volvox carteri reveal insights into evolution and origins of small RNA systems in plants. BMC Genomics 17, 853 (2016).

35. Fuchs, R. T., Sun, Z., Zhuang, F. & Robb, G. B. Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. PLoS One 10, e0126049 (2015).

36. Jackson, T. J., Spriggs, R. V., Burgoyne, N. J., Jones, C. & Willis, A. E. Evaluating bias-reducing protocols for RNA sequencing library preparation. BMC Genomics 15, 569 (2014).

37. Alon, S. *et al*. Barcoding bias in high-throughput multiplex sequencing of miRNA. Genome Res 21, 1506–1511 (2011).

38. Hafner, M. *et al*. RNA-ligase-dependent biases in miRNA representation in deep-sequenced small RNA cDNA libraries. RNA 17, 1697–1712 (2011).

39. Molnar, A. *et al*. Highly specific gene silencing by artificial microRNAs in the unicellular alga Chlamydomonas reinhardtii. Plant J
58, 165–174 (2009).

40. Zhao, T., Wang, W., Bai, X. & Qi, Y. Gene silencing by artificial microRNAs in Chlamydomonas. Plant J 58, 157–164 (2009).

41. Hausser, J. & Zavolan, M. Identification and consequences of miRNA-target interactions-beyond repression of gene expression. Nat Rev Genet 15, 599–612 (2014).

42. Liang, H., Zhang, J., Zen, K., Zhang, C. Y. & Chen, X. Nuclear microRNAs and their unconventional role in regulating non-coding RNAs. Protein Cell 4, 325–330 (2013).

43. Carroll, A. P., Tran, N., Tooney, P. A. & Cairns, M. J. Alternative mRNA fates identified in microRNA-associated transcriptome analysis. BMC Genomics 13, 561 (2012).

44. Chung, B. Y. W., Deery, M. J., Groen, A. J., Howard, J. & Baulcombe, D. C. mRNA turnover through CDS-targeting is the primary role of miRNA in the green alga Chlamydomonas. bioRxiv doi:10.1101/088807 (2016).

45. Herron, M. D., Hackett, J. D., Aylward, F. O. & Michod, R. E. Triassic origin and early radiation of multicellular volvocine algae. Proc Natl Acad Sci USA 106, 3254–3258 (2009).

46. Hertel, J. & Stadler, P. F. The Expansion of Animal MicroRNA Families Revisited. Life (Basel) 5, 905–920 (2015).

47. Axtell, M. J., Westholm, J. O. & Lai, E. C. Vive la différence: biogenesis and evolution of microRNAs in plants and animals. Genome Biol 12, 221 (2011).

48. Guo, W. *et al*. High-throughput sequencing and degradome analysis reveal neutral evolution of Cercis gigantea microRNAs and their targets. Planta 243, 83–95 (2016).

49. Mullokandov, G. *et al*. High-throughput assessment of microRNA activity and function using microRNA sensor and decoy libraries.
Nature Meth 9, 840–846 (2012).

50. Bosson, A. D., Zamudio, J. R. & Sharp, P. A. Endogenous miRNA and target concentrations determine susceptibility to potential ceRNA competition. Mol Cell 56, 347–359 (2014).

51. Denzler, R. *et al*. Impact of MicroRNA Levels, Target-Site Complementarity, and Cooperativity on Competing Endogenous RNA- Regulated Gene Expression. Mol Cell 64, 565–579 (2016).

52. Nozawa, M. *et al*. Evolutionary Transitions of MicroRNA-Target Pairs. Genome Biol Evol 8, 1621–1633 (2016).

53. Casas-Mollano, J. A. *et al*. Diversification of the core RNA interference machinery in Chlamydomonas reinhardtii and the role of DCL1 in transposon silencing. Genetics 179, 69–81 (2008).

54. Burroughs, A. M., Ando, Y. & Aravind, L. New perspectives on the diversification of the RNA interference system: insights from comparative genomics and small RNA sequencing. Wiley Interdisciplinary Reviews RNA 5, 141–181 (2014).

55. Tarver, J. E. *et al*. microRNAs and the evolution of complex multicellularity: identification of a large, diverse complement of microRNAs in the brown alga Ectocarpus. Nucleic Acids Res 43, 6384–6398 (2015).

56. Mattick, J. S. A new paradigm for developmental biology. J Exp Biol 210, 1526–1547 (2007).

57. Peterson, K. J., Dietrich, M. R. & McPeek, M. A. MicroRNAs and metazoan macroevolution: insights into canalization, complexity, and the Cambrian explosion. Bioessays 31, 736–747 (2009).

58. Avesson, L., Reimegård, J., Wagner, E. G. & Söderbom, F. MicroRNAs in Amoebozoa: deep sequencing of the small RNA population in the social amoeba Dictyostelium discoideum reveals developmentally regulated microRNAs. RNA 18, 1771–1782 (2012).

59. Sueoka, N. Mitotic Replication of Deoxyribonucleic Acid in Chlamydomonas Reinhardi. Proc Natl Acad Sci USA 46, 83–91 (1960).

60. van Dijk, K. *et al*. Monomethyl histone H3 lysine 4 as an epigenetic mark for silenced euchromatin in Chlamydomonas. Plant Cell
17, 2439–2453 (2005).

61. Ibrahim, F. *et al*. Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. Proc Natl Acad Sci USA 107, 3906–3911 (2010).

62. Merchant, S. S. *et al*. The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science 318, 245–250 (2007).

63. Lorenz, R. *et al*. Vienna RNA Package 2.0. Algorithms Mol Biol 6, 26 (2011).

64. Kruger, J. & Rehmsmeier, M. RNAhybrid: microRNA target prediction easy, fast and flexible. Nucleic Acids Res 34 (Web Server issue), W451–454 (2006).

65. Lopez, D., Casero, D., Cokus, S. J., Merchant, S. S. & Pellegrini, M. Algal Functional Annotation Tool: a web-based analysis suite to functionally interpret large gene lists using integrated annotation and expression data. BMC Bioinformatics 12, 282 (2011).

66. Goodstein, D. M. *et al*. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 40, D1178–1186 (2012). (Database issue).

67. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).

68. Schmollinger, S. *et al*. Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. Plant Cell 26, 1410–1435 (2014).

69. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA- Seq. Nature Meth 5, 621–628 (2008).

70. Anders, S. & Huber, W. Differential expression analysis for sequence count data. Genome Biol 11, R106 (2010).

71. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. J Mol Biol 215, 403–410 (1990).

CHAPTER 4

TARGETED GENE SILENCING BY RNA INTERFERENCE IN CHLAMYDOMONAS

Abstract

Small RNA-guided gene silencing is an evolutionarily conserved process that operates by a variety of molecular mechanisms and plays an essential role in developmental pathways and defense responses against genomic parasites in eukaryotes. Double-stranded RNA (dsRNA) triggered posttranscriptional gene silencing, termed RNA interference (RNAi), is also becoming a powerful tool for reverse genetics studies. Stable RNAi, induced by the expression of long dsRNAs or duplex small RNAs from genome-integrated transgenes, has been achieved in multiple organisms, including the green alga Chlamydomonas reinhardtii. However, the level of gene repression is often quite variable, depending on the type of construct, transgene copy number, site of integration, and target gene. Moreover, unintended transcripts partly complementary to a trigger dsRNA can also be silenced, making difficult the interpretation of observed phenotypes. To obviate some of these problems we have developed a tandem inverted repeat system that consistently induces cosilencing of a gene with a selectable RNAi-induced phenotype (encoding tryptophan synthase b-subunit) and any other (nonessential) gene of interest. In addition, to circumvent off-target effects, for each tested gene, RNAi lines are generated with at least two transgenes, homologous to

distinct and nonoverlapping sequences of the target transcript. A common phenotype among these independent RNAi strains is expected to result from suppression of expression of the gene of interest. We demonstrate this approach for the characterization of a gene of unknown function in Chlamydomonas, encoding a predicted exoribonuclease with weak similarity to 30hExo/ERI-1.

Introduction

RNA-mediated silencing is an evolutionarily conserved mechanism(s) by which double-stranded RNA (dsRNA) induces the inactivation of cognate sequences in eukaryotes, although dsRNA is also known to participate in the activation of gene expression (Carthew and Sontheimer, 2009; Cerutti and Casas-Mollano, 2006; Ghil- diyal and Zamore, 2009; Grimm, 2009). The role of long dsRNA in triggering repression was initially characterized in Caenorhabditis elegans and termed RNA interference (RNAi) (Fire *et al*., 1998). However, silencing phenomena had already been described in multiple species and the connection to dsRNA helped to unify several, apparently disparate, processes including posttranscriptional RNA degradation, translational repression, transcriptional gene silencing via heterochromatin formation and/or DNA methylation, DNA elimination, and meiotic silencing by unpaired DNA (Baulcombe, 2004; Cerutti and Casas-Mollano, 2006; Matzke and Birchler, 2005). In plants, animals, and the unicellular green alga Chlamydomonas reinhardtii, the RNAi machinery is also involved in the production of microRNAs (miRNAs), small RNAs resulting from the processing of genome encoded imperfect RNA hairpins, which play a role in developmental regulation

and other processes (Bartel, 2004; Carthew and Sontheimer, 2009; Casas-Mollano *et al*., 2008; Ghildiyal and Zamore, 2009; Molnar *et al*., 2007; Zhao *et al*., 2007).

The most extensively characterized dsRNA-mediated mechanism is transcript silencing triggered by small interfering RNAs (siRNAs). In this process, long dsRNAs are processed into siRNAs (20–25 nt in length) by an RNaseIII-like endonuclease, named Dicer (Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009). The siRNAs are then incorporated into a multiprotein complex, the RNA-induced silencing complex (RISC), where they associate with Argonaute (AGO) polypeptides (Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009). Recent evidence suggests that a siRNA duplex is loaded into RISC and then AGO cleaves one of the siRNA strands (the passenger strand) triggering its dissociation from the complex (Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009). Activated RISC then uses the remaining single-stranded siRNA as a guide to identify homologous RNAs, ultimately causing transcript degradation or translation repression (Carthew and Sontheimer, 2009; Ghildiyal and Zamore, 2009).

In slightly over a decade, RNAi has evolved from a fascinating biological phenomenon into a powerful experimental tool for reverse genetics studies and for inducing phenotypic changes for practical applications (Grimm, 2009; Hebert *et al*., 2008). RNAi can be used to suppress the expression of genes of interest by introduction into cells or whole organisms of exogenously synthesized dsRNAs/siRNAs or nonintegrative dsRNA/siRNA-producing viral vectors or plasmids (Grimm, 2009; Hebert *et al*., 2008; Waterhouse and Helliwell, 2003). However, the silencing effects of these exogenous nucleic acids are generally transient. Stable and heritable RNAi has also been

developed in a variety of eukaryotes, including Chlamydomonas (Grimm, 2009; Rohr *et al*., 2004; Schroda, 2006; Waterhouse and Helliwell, 2003). Common approaches rely on the production of antisense RNA or dsRNA by transcription from genome-integrated transgenes (Rohr *et al*., 2004; Schroda, 2006; Waterhouse and Helliwell, 2003). Yet, the level of silencing triggered by these transgenes is often variable and several (many) individual lines need to be molecularly characterized for suppression of a certain gene before potential phenotypic defects can be evaluated. More recently, transgenes mimicking the structure of miRNA precursors have been exploited to produce artificial miRNAs (amiRNAs) targeting genes of interest (Grimm, 2009; Molnar *et al*., 2009; Zeng *et al*., 2002; Zhao *et al*., 2009). In *C. reinhardtii*, the expression of amiRNAs appears to be more stable than that of long dsRNAs produced from inverted repeat (IR) transgenes but, in some cases, high levels of amiRNAs do not lead to target mRNA suppression (Molnar *et al*., 2009; Zhao *et al*., 2009). An additional issue with all RNAi approaches is that of target specificity since unintended transcripts, partly complementary to the trigger dsRNA, siRNA, or amiRNA, may also be silenced, a phenomenon termed “off-target” effects (Grimm, 2009; Kulkarni *et al*., 2006; Xu *et al*., 2006).

To overcome some of these problems, we have developed a tandem inverted repeat (TIR)–RNAi system that allows for the direct selection of effective transgenic RNAi lines and the simultaneous silencing of two unrelated genes (Figure 4.1A; Rohr *et al*., 2004). In Chlamydomonas, tryptophan synthase b-subunit (encoded by the MAA7 gene) converts the indole analog 5-fluoroindole (5-FI) into the toxic tryptophan analog 5-fluorotryptophan (Palombella and Dutcher, 1998). RNAi-mediated suppression of MAA7 leads to strains resistant to 5-FI (Rohr *et al*., 2004). Thus, selection for this RNAi-induced

phenotype permits the identification of transgenic lines showing effective interference of any (nonessential) gene co-targeted with MAA7, as part of the same dsRNA trigger synthesized from a TIR (Figure 4.1A; Rohr *et al*., 2004). In addition, for each gene of interest, RNAi lines are generated with at least two independent transgenes, homologous

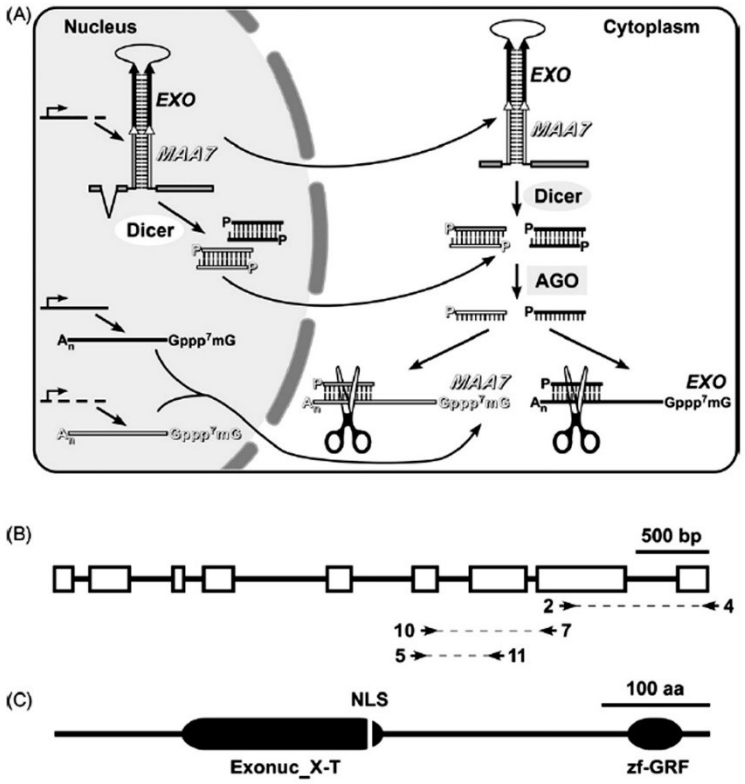

Figure 4.1. TIR–RNAi system for silencing of the EXO gene in Chlamydomonas. (A) Model of the TIR–RNAi system designed to target simultaneously the MAA7 and EXO transcripts. The subcellular localization of long dsRNA processing by Dicer in Chlamydomonas is presently hypothetical. Target mRNA cleavage is the depicted outcome of RNAi but the machinery can also operate by translation repression. (B) Diagram of the EXO precursor messenger RNA (excluding 50 and 30 untranslated regions), with exons indicated by white boxes. The annealing sites of primers used for RT-PCR amplification are shown underneath the exons. (C) Diagram depicting the domain architecture of the EXO protein. Exonuc_X-T, exonuclease, RNase T/DNA polymerase III domain; NLS, binary nuclear localization signal; zf-GRF, zinc finger motif named after three conserved residues in the center of the domain alignment.

to distinct and nonoverlapping sequences of the target transcript. A common phenotype among these strains is expected to result from suppression of expression of the desired gene rather than off-target effects. We demonstrate here this approach for the in vivo characterization of a predicted exonuclease (EXO) in Chlamydomonas.

Materials and Methods

5.1.1 Plasmid Construction

For generating an IR transgene targeting the EXO gene (protein ID 407657; http:// genomeportal.jgi-psf.org/Chlre4/Chlre4.home.html), a 580-bp fragment corresponding to the 30 end of its coding sequence was amplified by reverse transcription polymerase chain reaction (RT-PCR), using as template total RNA isolated from the CC-124 Chlamydomonas strain, with primers Exo-2 (5'-AAGCAGCCATGGTGGCAAG-3') and Exo-4 (5'-TACACACAAACACACGCTGACG-3') (Figure 4.1B). This PCR product was digested with SmaI/MluI to generate a 435-bp fragment that was cloned by blunt-end ligation into the MluI site of a pSTBlue-1 vector (Novagen, Madison, WI, USA), containing a 200-bp DNA spacer previously inserted into the EcoRV site (Rohr *et al*., 2004). Another copy of the 435-bp SmaI/MluI fragment was then inserted by blunt-end ligation into the HincII site of the generated vector. This way, two identical EXO segments were ligated flanking the DNA spacer. The orientation of the cloned fragments was examined by digestion with PvuII, which cuts only once within the 435-bp EXO sequence, and a plasmid containing the inserts in inverted orientation was verified by DNA sequencing. Standard protocols were used for RT-PCR, restriction enzyme

digestion, generation of blunt ends with T4 DNA polymerase, and DNA ligation (Sambrook and Russell, 2001).

The EXO3 IR cassette was then excised by digestion with PstI/HindIII and inserted, by blunt-end ligation, into the EcoRI sites of the MAA7/X IR vector (Rohr *et al*., 2004) to generate the MAA7/EXO3 IR transgene. This construct contains a TIR with homology to the MAA7 and the EXO genes so that transcription leads to the production of a hairpin-loop-shaped RNA that can be processed by the RNAi machinery into siRNAs targeting both MAA7 and EXO transcripts (Figure 4.1A). The MAA7/X IR vector contains, in addition to a MAA7 IR, an engineered aphVIII gene, encoding aminoglycoside 30-phosphotransferase (Sizova *et al*., 2001), confer- ring resistance to paromomycin for selection purposes (Rohr *et al*., 2004). A similar strategy was used to construct the MAA7/EXO5 IR transgene, which is homologous to the middle region of the EXO coding sequence. Briefly, a 490-bp fragment was amplified by RT-PCR with primers Exo-10 (50-CCTGTTCACACAGCGCTACAA- 30) and Exo-7 (50-ATCGCTCCTGCTCCTGCTC-30) (Figure 4.1B). This PCR product was then cloned, by blunt-end ligation, in forward and reverse orientations into the MluI and HincII sites of the pSTBlue-1 vector, flanking the 200-bp DNA spacer (Rohr *et al*., 2004). The EXO5 IR cassette was then excised by digestion with PstI/HindIII and cloned, by blunt-end ligation, into the EcoRI sites of the MAA7/X IR vector (Rohr *et al*., 2004) to generate the MAA7/EXO5 IR transgene. The EXO sequences used to build the IR constructs were chosen because they are distinct and display very low homology (no contiguous stretch longer than 19 nt) to nontarget genes in BLAST searches of the Chlamydomonas genome

(http://genomeportal.jgi- psf.org/Chlre4/Chlre4.home.html), in order to minimize the possibility of off-target effects (Kulkarni *et al*., 2006).

5.1.2 Culture Conditions, Cell Transformation, and Selection Procedures

Chlamydomonas reinhardtii cells were grown in Tris–acetate–phosphate (TAP) medium (Harris, 1989). Plasmid DNA linearized with SacI, cleaving upstream from the transgenic promoter, was used in all transformation experiments. The wild type strain CC-124 (mt–) (Harris, 1989) was transformed by the glass beads procedure (Kindle, 1990) and allowed to recover for 2 days, to permit induction of RNAi, before plating under selective conditions. MAA7/EXO3 IR or MAA7/EXO5 IR transformants were selected on TAP medium containing 10 μg/ml paromomycin and 7.5 μM 5-FI (Sigma, St. Louis, MO, USA). Plates were incubated under dim lights (~50 μmol/m2/s photosynthetically active radiation) to lessen the light-induced degradation of 5-FI (Palombella and Dutcher, 1998). In our previous system, the selective medium was supplemented with 1.5 mM L-tryptophan (Rohr *et al*., 2004). However, addition of this amino acid was found to be unnecessary for survival or normal growth of the Chlamydomonas transformants. As a negative control, CC-124 was also transformed with the MAA7/X IR vector (Rohr *et al*., 2004), which only contains the IR transgene designed to induce RNAi of MAA7. All isolated transgenic strains were kept under constant selective pressure to circumvent any potential loss of activity due to silencing of the integrated IR transgenes. For phenotypic analyses, cells grown to logarithmic phase in liquid TAP medium were serially diluted, spotted on plates of the appropriate media

(TAP or TAP containing 7.5 μM 5-FI), and incubated for 10–15 days under dim lights (Zhang *et al*., 2002).

5.1.3 RNA Analyses

Total cell RNA was purified with TRI Reagent (Molecular Research Center, Cincinnati, OH, USA), according to the manufacturer's instructions. The isolated RNA was separated by agarose/formaldehyde gel electrophoresis, blotted onto nylon membranes, and hybridized with 32P-labeled probes specific for ACT1 (encoding actin) or the coding sequence of histone H2A (Sambrook and Russell, 2001; Wu-Scharf *et al*., 2000; Zhang *et al*., 2002).

5.1.4 Reverse Transcriptase-PCR Analyses

TRI Reagent isolated total RNA was treated with DNase I (RNase-free, Ambion, Austin, TX, USA) to remove contaminating DNA and reverse transcription reactions were carried out as previously described (Carninci *et al*., 1998). The synthesized cDNA was then used as a template in standard PCR reactions (Sambrook and Russell, 2001; Wu- Scharf *et al*., 2000). The numbers of cycles showing a linear relationship between input RNA and the final product were determined in preliminary experiments. Primers were designed to match exonic sequences flanking one or more introns, to distinguish con- taminating PCR products generated by the amplification of any remaining DNA. Controls also included the use as template of reactions without RT and verification of PCR products by hybridization with specific probes (data not shown). The PCR conditions for amplification of ACT1 were 23 cycles at 93°C for 30 s, at 55°C for 30 s,

and at 71°C for 60 s. The EXO PCR product was amplified by 32 cycles under the same conditions. Aliquots (5-µl) of each RT-PCR reaction were resolved on 2% agarose gels and visualized by ethidium bromide staining. The primer sequences were as follows: Exo-5, 5'-AACCTGAAGCGCCTGTTCAC-3', EXO-11, 5'-GCTGCTCCCTGAAGCT GTCT-3'; ACT-cod-F (5'-GACATCCGCAAGGACCTCTAC-3'), and ACT-cod-R (5'-GATCCACATTTGCTGGAAGGT-3').

Results and Discussion

The predicted Chlamydomonas EXO protein belongs to the DnaQ-H superfamily of 3' to 5' exonucleases, with a DEDDh motif that is characteristic of the 3'hExo/ERI-1 subfamily (Kupsco *et al*., 2006). The 3'hExo exoribonuclease was initially identified as a candidate regulator of histone mRNA degradation in mammalian cells (Dominski *et al*., 2003), but its in vivo role is not clear since the decay of histone transcripts involves similar pathways to those required for the degradation of poly(A) mRNAs after deadenylation (Marzluff *et al*., 2008). Intriguingly, the closest *C. elegans* and *Schizosaccharomyces pombe* homologs of 3'hExo, named ERI-1, have been implicated as negative regulators of RNAi and in 5.8S rRNA processing (Gabel and Ruvkun, 2008; Iida *et al*., 2006; Kennedy *et al*., 2004). Another member of this subfamily, Drosophila melanogaster Snipper, efficiently degrades structured, dsRNA, and DNA substrates as long as there exists a 3' overhang of a few nucleotides to initiate decay (Kupsco *et al*., 2006). However, the in vivo function of Snipper remains unknown since Snp mutant flies are viable and display no obvious phenotypic abnormalities (Kupsco *et al*., 2006). The

human Snipper homolog, named exonuclease domain containing 1 (ExoD1), is produced in two isoforms (Gene ID 112479): a shorter one similar in structure to Snipper and a longer one that includes, in addition to the exonuclease domain, a zinc finger of the GRF (Glycine-Arginine-Phenylalanine) type. Thus, exonucleases belonging to the 3'hExo/ERI-1 subfamily appear to show activity toward a diverse array of substrates and play a variety of roles.

Chlamydomonas EXO is most similar in sequence and domain structure (Figure 4.1C) to the long human ExoD1 isoform and lacks a nucleic acid-binding SAP (after SAF-A/B, Acinus and PIAS) motif, typical of ERI-1 proteins. We used the TIR–RNAi system (Figure 4.1A) as a useful tool to gain insight into the function of the uncharacterized EXO protein. In addition, we were interested in examining the possibility that EXO might act as a negative regulator of RNAi in Chlamydomonas and its suppression might, therefore, lead to strains with enhanced RNAi. We constructed two tandem IR trans- genes targeting simultaneously the 30 untranslated region of the MAA7 mRNA as well as distinct, nonoverlapping regions of the EXO transcript: the MAA7/EXO3 IR transgene produces dsRNA homologous to the 3' end of the EXO coding sequence (Figure 4.1B, region flanked by primers 2 and 4) and the MAA7/EXO5 IR transgene generates dsRNA matching the middle of the EXO coding sequence (Figure 4.1B, region flanked by primers 10 and 7). Chlamydomonas cells resistant to 5-FI were isolated by transformation with either the MAA7/EXO3 IR transgene (Figure 4.2A) or the MAA7/ EXO5 IR transgene (data not shown). Interestingly, transgenic lines transformed with these TIRs grew slower than the wild-type CC-124 on solid or in liquid TAP medium (Figure 4.2A and data not shown). By contrast, Chlamydomonas strains

containing an IR transgene targeting exclusively the MAA7 gene can also survive on medium containing 5-FI but are not affected in their growth rate (Figure 4.2A, Maa7-IR5; Rohr *et al*., 2004). The transgenic lines transformed with the MAA7/EXO IR containing plasmids and resistant to 5-FI also showed reduced levels of the EXO transcript in semiquantitative RT-PCR assays (Figure 4.2B), consistent with RNAi-mediated suppression of both the MAA7 and EXO genes. This observation applied to strains generated with the MAA7/ EXO3 IR vector (Figure 4.2B, Exo3-IR5 and Exo3-IR8) as well as those generated with the MAA7/EXO5 IR transgene (Figure 4.2B, Exo5-IR7 and Exo5-IR11). However, the degree of EXO gene suppression varied among the different transgenic lines. In contrast, the Maa7-IR5 strain, containing the IR transgene designed to downregulate exclusively MAA7, did not display any decrease in EXO mRNA levels when compared with the wild type (Figure 4.2B). Thus, the TIR-RNAi system, as previously demonstrated for several unrelated genes (Rohr *et al*., 2004), allowed the recovery of effective EXO-suppressed RNAi strains by selection for MAA7 silencing in medium containing 5-FI. The ERI-1 exonuclease has been implicated in the degradation of siRNA duplexes with two- nucleotide 3'overhangs, reducing the efficiency of RNAi (Kennedy *et al*., 2004), and in endogenous small RNA pathways in *C. elegans* (Lee *et al*., 2006). Thus, we tested whether Chlamydomonas EXO might act as an RNAi regulator in the transgenic lines.

However, the EXO-suppressed RNAi strains did not show any change in the levels of several endogenous small RNAs when compared with the wild-type CC-124 (data not shown). The mammalian homolog of ERI-1, 3'hExo has been proposed to play a role, although presently undefined, in histone mRNA metabolism (Dominski *et al*.,

2003; Yang *et al*., 2009). Interestingly, when we examined the Chlamydomonas lines containing the MAA7/EXO IR transgenes for histone H2A transcript amounts they displayed a significant reduction in comparison with the control strains (Figure 4.2C). Most core histone genes in Chlamydomonas, similarly to those in metazoans, have a highly conserved palindromic sequence (that can fold into a stem–loop structure) at their

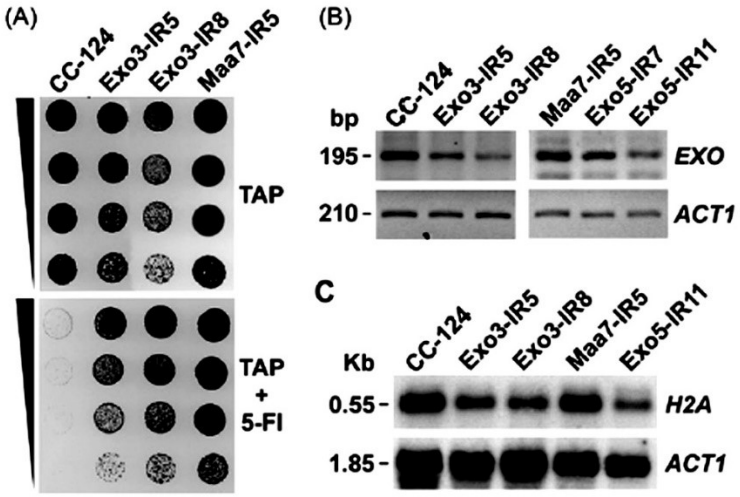

Figure 4.2. Silencing of the EXO gene by tandem IR transgenes in transformants selected on 5-FI-containing medium. (A) Growth and survival of the indicated strains on TAP medium without (upper panel) or with (lower panel) 7.5μM of 5-FI. CC-124, untransformed wild-type strain; Exo3-IR5 and Exo3-IR8, CC-124 transformed with a tandem IR transgene inducing simultaneous RNAi of the MAA7 (encoding tryptophan synthase b-subunit) and the EXO genes; Maa7-IR5, CC-124 transformed with an IR transgene designed to induce RNAi of MAA7. (B) Semiquantitative RT-PCR analysis of EXO gene expression in the indicated strains. Amplification of ACT1 (encoding actin) transcripts is shown as an input control. Reactions using RNA not treated with reverse transcriptase as the template were employed as a negative control (data not shown). Exo5-IR7 and Exo5-IR11, CC-124 transformed with a tandem IR transgene inducing simultaneous RNAi of the MAA7 and the EXO genes but targeting a different region of the EXO transcript than the MAA7/EXO3 IR (see text for details). (C) Northern blot analysis of total cell RNA sequentially hybridized with 32P-labeled PCR products corresponding to the coding sequence of histone H2A (upper panel), to evaluate the degree of mRNA reduction, or the coding sequence of ACT1 (lower panel), to assess the amount of sample loaded per lane.

3'end, within a short distance of the stop codon (Fabry *et al*., 1995). Replication-dependent histone mRNAs lack a poly(A) tail in metazoans and their 3' end is formed by endonucleolytic cleavage very near the 30 stem–loop sequence (Marzluff *et al*., 2008). A single protein, the stem-loop-binding protein (SLBP), binds to this conserved hairpin structure and participates in many aspects of histone mRNA metabolism, including 3'end processing (Marzluff *et al*., 2008). Knockdown of SLBP results in reduced histone transcript levels, nuclear retention of histone mRNA, and defects in mammalian cell division (Sullivan *et al*., 2009). These phenotypes are very similar to those of the Chlamydomonas EXO RNAi strains, although we did not examine the possible nuclear retention of histone transcripts. Nonetheless, it is tempting to speculate that the EXO protein, like mammalian SLBP, may be required for the maturation of the 3'end of histone mRNAs in Chlamydomonas. In its absence, improperly processed histone transcripts may be unstable and subject to degradation by RNA quality control mechanisms (Doma and Parker, 2007).

A number of approaches have been successfully used for RNA-mediated suppression of gene expression in Chlamydomonas, including genome-integrated transgenes generating long dsRNAs or amiRNAs (Molnar *et al*., 2009; Rohr *et al*., 2004; Schroda, 2006; Zhao *et al*., 2009). IR transgenes producing long dsRNAs have the disadvantage that they may be prone to self-silencing at the transcriptional level (Rohr *et al*., 2004; Yamasaki *et al*., 2008), although this may be partly avoided by expression from an inducible promoter (Koblenz and Lechtreck, 2005). In addition, long dsRNAs can be processed into a complex and heterogeneous population of siRNAs with the possibility that some of these small RNAs will match unintended transcripts and cause off-target

effects (Grimm, 2009; Kulkarni *et al*., 2006; Xu *et al*., 2006). In contrast, amiRNA transgenes appear to be more stable in their expression and, since each amiRNA precursor gives rise to a single small RNA species, they can be optimized to avoid off-target effects (Molnar *et al*., 2009; Zhao *et al*., 2009). However, there are also disadvantages with the amiRNA approach since many aspects of small RNA processing, regulation, and miRNA action remain unknown in Chlamydomonas (Casas-Mollano *et al*., 2008). Thus, designed amiRNAs may not be processed as intended or the unique site where an amiRNA would associate on a target transcript may not be accessible, for instance, because of secondary structure formation (Grimm, 2009; Molnar *et al*., 2009). This may result in the accumulation of ineffective amiRNAs with certain transgenes and require careful design of more than one amiRNA construct for any given target gene (Molnar *et al*., 2009).

The TIR–RNAi system obviates some of these issues. In addition, a recently reported PCR strategy can greatly simplify the construction of IR transgenes (Pawloski *et al*., 2005). Transcriptional self-silencing and other problems associated with genomic transgene integration (Rohr *et al*., 2004) are avoided by directly selecting for an RNAi-induced phenotype. Moreover, this also allows the recovery of effective RNAi strains against any (nonessential) gene cotargeted with the selectable MAA7 marker.

Conceptually, a similar result can be achieved with dimeric amiRNA precursors (Molnar *et al*., 2009; Zhao *et al*., 2009). However, the production of long dsRNAs simplifies the design of RNAi transgenes because of the increased probability, relative to a single amiRNA, that at least some of the processed siRNAs will be assembled into a functional RISC and will have accessibility to the target transcript. The greater risk of

off-target effects when using long dsRNA can be circumvented by silencing each gene of interest with at least two transgenes, homologous to distinct and nonoverlapping sequences of the transcript. The finding of similar phenotypes in several independent RNAi strains generated with different transgenes increases confidence that any phenotypic abnormalities are the consequence of downregulation of the intended gene. Moreover, there should be a clear correlation between target transcript and/or protein depletion and phenotypic severity. As demonstrated here for Chlamydomonas EXO, this approach can be used for the characterization of genes of unknown function.

Summary

RNAi is rapidly becoming a standard method for experimental and practical gene silencing. In the unicellular green alga *Chlamydomonas reinhardtii*, work from many groups has generated a number of tools for effective RNAi, including the TIR system described here. Both IR transgenes, producing long dsRNA, as well as amiRNA transgenes, generating unique small RNA molecules, have been successfully used for targeted gene silencing. However, potential drawbacks to this promising technology have also become apparent, such as a possible lack of specificity resulting in off-target effects. Moreover, relatively little is known about the role of the RNAi machinery in gene regulation and other processes in Chlamydomonas and the consequences that overexpression of transgenic RNAi triggers may have on these endogenous mechanisms. This emphasizes the need for careful design of RNAi experiments, including appropriate negative controls and the use of multiple amiRNA or long dsRNA species for each target

gene. When- ever possible, data from RNAi-mediated knockdown should also be validated with alternative approaches. In addition, although throughout this chapter we have accentuated reduction in target transcript level as the expected outcome of RNAi, recent evidence suggests that RNAi may also operate by translational repression in Chlamydomonas requiring the assessment of protein amounts in transgenic RNAi lines. Ultimately, despite its limitations, RNAi provides a powerful method for uncovering gene function in organisms where gene targeting by homologous recombination is not practical.

References

Bartel, D.P. (2004). MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116, 282–297. Baulcombe, D. (2004). RNA silencing in plants. Nature 431, 356–363.

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. (1998). Thermostabilization and thermoactivation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. USA 95, 520–524.

Carthew, R.W., and Sontheimer, E.J. (2009). Silence from within: Endogenous siRNAs and miRNAs. Cell 136, 642–655.

Casas-Mollano, J.A., Rohr, J., Kim, E.J., Balassa, E., van Dijk, K., and Cerutti, H. (2008). Diversification of the core RNA interference machinery in Chlamydomonas reinhardtii and the role of DCL1 in transposon silencing. Genetics 179, 69 - 81.

Cerutti, H., and Casas-Mollano, J.A. (2006). On the origin and functions of RNA-mediated silencing: From protists to man. Curr. Genet. 50, 81–99.

Doma, M.K., and Parker, R. (2007). RNA quality control in eukaryotes. Cell 131, 660–668.

Dominski, Z., Yang, X.-C., Kaygun, H., Dadlez, M., and Marzluff, W.F. (2003). A 30 exonuclease that specifically interacts with the 30 end of histone mRNA. Mol. Cell 12, 295–305.

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.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.

Gabel, H.W., and Ruvkun, G. (2008). The exonuclease ERI-1 has a conserved dual role in 5.8S rRNA processing and RNAi. Nat. Struct. Mol. Biol. 15, 531–533.

Ghildiyal, M., and Zamore, P.D. (2009). Small silencing RNAs: An expanding universe. Nat. Rev. Genet. 10, 94–108.

Grimm, D. (2009). Small silencing RNAs: State-of-the-art. Adv. Drug Deliv. Rev. 61, 672–703.

Harris, E.H. (1989). “The Chlamydomonas Sourcebook—A Comprehensive Guide to Biology and Labora- tory Use.” Academic Press, San Diego.

Hebert, C.G., Valdes, J.J., and Bentley, W.F. (2008). Beyond silencing—engineering applications of RNA interference and antisense technology for altering cellular phenotype. Curr. Opin. Biotechnol. 19, 500–505.

Iida, T., Kawaguchi, R., and Nakayama, J. (2006). Conserved ribonuclease, Eri1, negatively regulates heterochromatin assembly in fission yeast. Curr. Biol. 16, 1459–1464.

Kennedy, S., Wang, D., and Ruvkun, G. (2004). A conserved siRNA-degrading RNase negatively regulates RNA interference in C. elegans. Nature 427, 645–649.

Kindle, K.L. (1990). High frequency nuclear transformation of Chlamydomonas reinhardtii. Proc. Natl. Acad. Sci. USA 87, 1228–1232.

Koblenz, B., and Lechtreck, K.F. (2005). The NIT1 promoter allows inducible and reversible silencing of centrin in Chlamydomonas reinhardtii. Eukaryot. Cell 4, 1959–1962.

Kulkarni, M.M., Booker, M., Silver, S.J., Friedman, A., Hong, P., Perrimon, N., and Mathey-Prevot, B. (2006). Evidence of off-target effects associated with long dsRNAs in Drosophila melanogaster cell-based assays. Nat. Methods 3, 833–838.

Kupsco, J.M., Wu, M.-J., Marzluff, W.F., Thapar, R., and Duronio, R.J. (2006). Genetic and biochemical characterization of Drosophila Snipper: A promiscuous member of the metazoan 30 hExo/ERI-1 family of 3’ to 5’exonucleases. RNA 12, 2103–2117.

Lee, R.C., Hammell, C.M., and Ambros, V. (2006). Interacting endogenous and exogenous RNAi pathways in Caenorhabditis elegans. RNA 12, 589–597.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: Life without a poly(A) tail. Nat. Rev. Genet. 9, 843–854.

Matzke, M.A., and Birchler, J.A. (2005). RNAi-mediated pathways in the nucleus. Nat. Rev. Genet. 6, 24–35.

Molnar, A., Schwach, F., Studholme, D.J., Thuenemann, E.C., and Baulcombe, D.C. (2007). miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447, 1126–1129.

Molnar, A., Bassett, A., Thuenemann, E., Schwach, F., Karkare, S., Ossowski, S., Weigel, D., and Baul- combe, D. (2009). Highly specific gene silencing by artificial microRNAs in the unicellular alga Chla- mydomonas reinhardtii. Plant J. 58, 165–174.

Palombella, A.L., and Dutcher, S.K. (1998). Identification of the gene encoding the tryptophan synthase b- subunit from Chlamydomonas reinhardtii. Plant Physiol. 117, 455–464.

Pawloski, L.C., Deal, R.B., McKinney, E.C., Burgos-Rivera, B., and Meagher, R.B. (2005). Inverted repeat PCR for the rapid assembly of constructs to induce RNA interference. Plant Cell Physiol. 46, 1872–1878.

Rohr, J., Sarkar, N., Balenger, S., Jeong, B.-R., and Cerutti, H. (2004). Tandem inverted repeat system for selection of effective transgenic RNAi strains in Chlamydomonas. Plant J. 40, 611–621.

Sambrook, J., and Russell, D.W. (2001). “Molecular Cloning—A Laboratory Manual.” Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schroda, M. (2006). RNA silencing in Chlamydomonas: Mechanisms and tools. Curr. Genet. 49, 69–84.

Sizova, I., Fuhrmann, M., and Hegemann, P.A. (2001). Streptomyces rimosus aph VIII gene coding for a new type phosphotransferase provides stable antibiotic resistance to Chlamydomonas reinhardtii. Gene 277, 221–229.

Sullivan, K.D., Mullen, T.E., Marzluff, W.F., and Wagner, E.J. (2009). Knockdown of SLBP results in nuclear retention of histone mRNA. RNA 15, 459–472.

Waterhouse, P.M., and Helliwell, C.A. (2003). Exploring plant genomes by RNA-induced gene silencing. Nat. Rev. Genet. 4, 29–38.

Wu-Scharf, D., Jeong, B.-r., Zhang, C., and Cerutti, H. (2000). Transgene and transposon silencing in Chlamydomonas reinhardtii by a DEAH-box RNA helicase. Science 290, 1159–1162.

Xu, P., Zhang, Y., Kang, L., Roossinck, M.J., and Mysore, K.S. (2006). Computational estimation and experimental verification of off-target silencing during posttranscriptional gene silencing in plants. Plant Physiol. 142, 429–440.

Yamasaki, T., Miyasaka, H., and Ohama, T. (2008). Unstable RNAi effects through epigenetic silencing of an inverted repeat transgene in Chlamydomonas reinhardtii. Genetics 180, 1927–1944.

Yang, X.C., Torres, M.P., Marzluff, W.F., and Dominski, Z. (2009). Three proteins of the U7-specific Sm ring function as the molecular ruler to determine the site of 30 -end processing in mammalian histone pre- mRNA. Mol. Cell Biol. 29, 4045–4056.

Zeng, T., Wagner, E.J., and Cullen, B.R. (2002). Both natural and designed micro RNAs can inhibit the expression of cognate mRNAs when expressed in human cells. Mol. Cell 9, 1327–1333.

Zhang, C., Wu-Scharf, D., Jeong, B.-R., and Cerutti, H. (2002). A WD40-repeat containing protein, similar to a fungal co-repressor, is required for transcriptional gene silencing in Chlamydomonas. Plant J. 31, 25–36.

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G.J., Wang, X.J., and Qi, Y. (2007). A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. Genes Dev. 21, 1190–1203.

Zhao, T., Wang, W., Bai, X., and Qi, Y. (2009). Gene silencing by artificial microRNAs in Chlamydomonas. Plant J. 58, 157–164.

CHAPTER 5

THE CHARACTERIZATION OF A CCR4-LIKE PROTEIN IN CHLAMYDOMONAS REINHARDTTI

Abstract

Carbon catabolite repressor 4 (CCR4) is a component of the CCR4/NOT complex. The CCR4 protein has been identified in *Saccharomyces cerevisiae* as a major cytoplasmic deadenylase. CCR4 contains an Exonuclease-Endonuclease-Phosphatase (EEP) domain that promotes shortening of the poly(A) tail of eukaryotic mRNAs and stimulates their degradation. Interestingly, recent studies in animals have suggested that CCR4 also participates in sRNA-mediated translational repression. However, the precise mechanisms of translation inhibition by sRNAs are still unclear in animals and also virtually unexplored in plants.

Here, we report that a CCR4-like protein is also involved in sRNA-mediated gene silencing in the unicellular alga *Chlamydomonas reinhardtii*. We have isolated an RNAi-defective mutant (*Mut26*), which contains a deletion of the gene encoding the Chlamydomonas homolog of CCR4. *Mut26* shows alterations in translational repression triggered by both siRNAs and miRNAs. However, in a poly(A) tail assay, siRNA-repressed transcripts are not altered in their polyadenylation status within *Mut26*, indicating that siRNA-mediated translation inhibition seems to occur in a deadenylation-independent manner. Interestingly, a Flag-tagged CCR4 protein was found to associate

with translation initiation complexes and mRNA decay complexes. Our results are consistent with CCR4 having a role in sRNA-dependent translational repression in Chlamydomonas and imply deep evolutionary conservation of these mechanisms of sRNA-mediated gene silencing, likely present in the last common ancestor of plants and animals.

Introduction

MicroRNAs (miRNAs) are endogenous non-coding short RNAs (~21-23nts) that serve as posttranscriptional regulators of gene silencing (Ameres & Zamore, 2013; Nishimura & Fabian, 2016). Animal miRNAs recognize target mRNAs by imperfectly base pairing via a seed region (nucleotides 2-8 of the miRNA guide strand) with the 3' UTR of target mRNAs to negatively impact their protein production. Seed pairing often occurs with central mismatches that allow translational repression rather than target cleavage (D. P. Bartel, 2009; Wilczynska & Bushell, 2015). In contrast, plant miRNAs often trigger cleavage of target mRNAs perfectly complementary to their sequences (Ameres & Zamore, 2013; Yamasaki et al., 2013). However, increasing genetic evidence suggests that plant miRNAs regulate gene expression by both target cleavage and translational repression. In Arabidopsis, miRNA172 is highly complementary to APETALA2 (a floral homeotic gene) and regulates its target transcript primarily through translation inhibition (Chen, 2004). Current biochemical analyses have provided more conclusive evidence to support these findings of translational repression by plant miRNAs (Iwakawa & Tomari, 2013; Nishimura & Fabian, 2016; Yamasaki et al., 2013).

Therefore, small RNA-mediated translational repression is conserved in both plants and animals (Brodersen *et al*., 2008; Iwakawa & Tomari, 2013; Jonas & Izaurralde, 2015; Wilczynska & Bushell, 2015).

Small RNAs (sRNAs) bind to Argonaute (Ago) proteins and serve as guides for the polypeptide to target the complementary sites on target transcripts. Ago protein is a crucial component, with slicer activity, in the small RNA-mediated regulatory pathway. This protein consists of four main domains, including N-terminal, PAZ (recognizing the 2-nt 3'overhang of sRNAs), MID, and PIWI (RNase H enzyme) domains (Miyoshi, Ito, Murakami, & Uchiumi, 2016; Song et al., 2003). Structural biology reporting on the prokaryotic Ago protein from *Pyrococcus furiosus* and *Aquifex aeolicus* elucidated the role of the PIWI domain as a site-specific RNA endonuclease (Song, Smith, Hannon, & Joshua-Tor, 2004; Yesuan *et al*., 2005). Hence, Ago proteins act known to play a primary role in the cleavage of target transcripts using their intrinsic endonuclease activity. However, as already mentioned in animals, miRNAs associated with Ago predominantly trigger translational repression rather than target cleavage, resulting in lost its catalytic activity.

Animal miRNAs reduce protein production by first repressing the initiation step of translation (Nishimura & Fabian, 2016). Additionally, miRNAs could also promote mRNA destabilization by recruiting the deadenylase complex (CCR4/NOT) through a GW182 protein that interacts with the miRNA-associated Ago complex (Iwakawa & Tomari, 2015). Despite extensive work on how miRNAs bind to mRNAs to trigger translational repression and mRNA decay (Iwakawa & Tomari, 2015; Nishimura & Fabian, 2016), their precise mode of action remains controversial. The CCR4 protein

being a component of the CCR4/NOT complex, contains an Exonuclease-Endonuclease-Phosphatase (EEP) domain that promotes shortening of the 3'-poly(A) tail of eukaryotic mRNAs and stimulates their degradation (Collart & Panasenko, 2012). Contradictorily, modern work has also shown that CCR4 participates in miRNA-mediate translational repression (Bhandari, Raisch, Weichenrieder, Jonas, & Izaurralde, 2014; Huntzinger & Izaurralde, 2011; Iwakawa & Tomari, 2015). Therefore, this project aims to uncover the role of the CCR4-like protein in sRNA-mediated translational repression in *Chlamydomonas reinhardtii*.

Results

1. A CCR4-like deletion mutant is defective in RNAi in Chlamydomonas

In *C. reinhardtii*, RNA interference (RNAi) has been used to understand the function of genes associated with photosynthesis, Carbon dioxide Concentrating Mechanisms (CCM), and small RNA biogenesis (Duanmua, Millerb, Horkenb, Weeksb, & Spalding, 2009; Ibrahim et al., 2010; Oey et al., 2013; Pollock, Prout-Jr, Godfrey, Lemaire, & Moroney, 2004). By random insertional mutagenesis, we have generated a Chlamydomonas mutant defective in RNAi and, thus, unable to silence the endogenous gene Maa7 (encoding tryptophan synthase β-subunit). One such mutant, *Mut26*, displayed a deletion of the gene encoding the *C. reinhardtii* homolog of the CCR4 protein. While a transgenic strain, Maa7-IR 44s, was able to grow in 7.0 μM of the indole analog 5-fluoroindole (5-FI), the mutant strain was resistant to this drug (Figure 5.1.A). Maa7-IR 44s contains an inverted repeat transgene, targeted for silencing the Maa7 gene

that is required to produce tryptophan synthase, which converts 5-FI into the toxic tryptophan analog 5-fluoro tryptophan. As previously described (Ma et al., 2013), in Maa7 IR 44s, the MAA7 repression occurs predominantly at the transcriptional level.

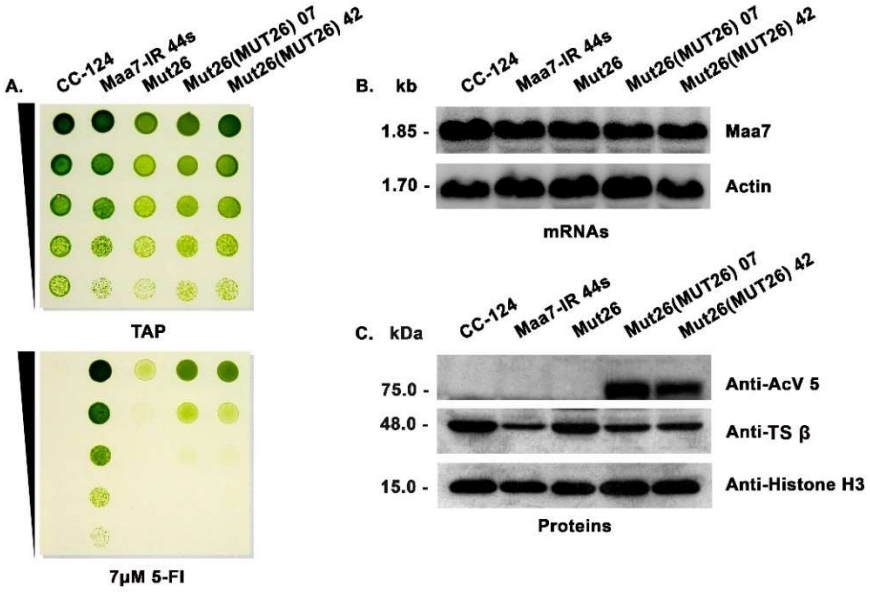

Figure 5.1. The MAA7 protein is increased in the RNAi defective *Mut26*, containing a deletion of the CCR4 gene. (A) Phenotypic analysis of the indicated strains grown in TAP medium with or without 7.0 μM of 5-Fluoroindole (5-FI). CC-124, wild-type; Maa7-IR 44s, a strain containing an inverted repeat transgene targeting Maa7; *Mut26*, a strain deleted for the *C. reinhardtii* CCR4 homolog; and Mut26(MUT26) 07 and Mut26(MUT26) 42, independent lines complemented with a full-length cDNA of CCR4. (B) Detection of Maa7 transcript levels by Northern blotting. Actin serves as a loading control. (C) Immunoblot detection of endogenous and/or transgenic proteins. The abundance of AcV5-tagged proteins and tryptophan synthase beta subunits (MAA7; TS-β).

Consistent with this observation, all strains had similar levels of Maa7 mRNAs (Figure 5.1.B). In contrast, *Mut26* displayed enhanced levels (very similar to those in the wild type) of the tryptophan synthase β subunit (Figure 5.1.C). Transformation of *Mut26* with a transgene containing the full-length of CCR4 coding sequence fused to a FLAG-CBP-AcV5 tag at its N terminus resulted in strains [Mut26(MUT26) 07 and Mut26(MUT26)

42] with partial to full reversion of the mutant phenotypes (Figure 5.1). These results suggested that Chlamydomonas CCR4 is involved in sRNA-mediated translational repression.

2. *Mut26* is also defective in translational repression of the endogenous Chlamydomonas miRNA target

To examine the putative role of the Chlamydomonas CCR4 homolog on

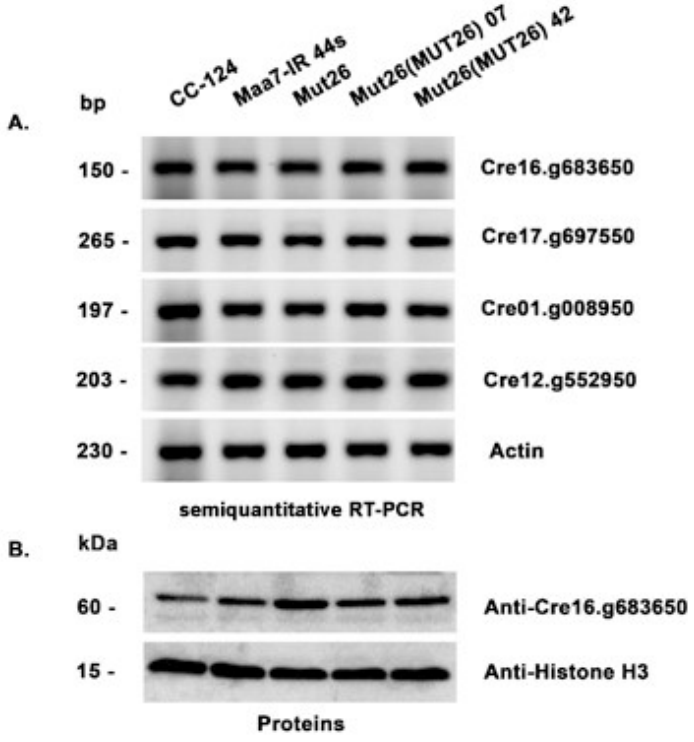

Figure 5.2. Analysis of miRNA target transcript and protein abundance. (A) qRT-PCR analyses of the transcript levels of miRNA targets in the indicated strains. *Cre16.g683650* is a target of miR C-mediated translational repression. *Cre17.g697550* and *Cre01.g008950* (encoding proteins of unknown function) are targets of miR B-mediated cleavage. *Cre12.g552950* (encoding a protein of unknown function) is a target of miR_t70-mediated cleavage. (B) Immunoblot detection of the *Cre16.g683650* protein (encoding a putative protein kinase). Histone H3 was used as a loading control.

endogenous gene regulation, we tested the abundance of three transcripts targeted for miRNA-mediated cleavage (*Cre17.g697550*, *Cre01.g008950*, and *Cre12.g552950* (Voshall et al., 2015; 2017) and one transcript targeted for miRNA-mediated translational repression (*Cre16.g683650*) (Voshall et al., 2015). The transcript levels for these four targets were similar to that in the wild type in all the strains examined (Figure 5.2.A), indicating that *Mut26* was not defective in miRNA-mediated transcript cleavage. However, the protein encoded by *Cre16.g683650* showed greater abundance in *Mut26*, consistent with a defect in siRNA-mediated translational repression (Figure 5.2.B). Thus, the deletion of CCR4 appears to disrupt the translational repression triggered by either siRNAs or miRNAs in Chlamydomonas, but has no consequence for miRNA-mediated degradation.

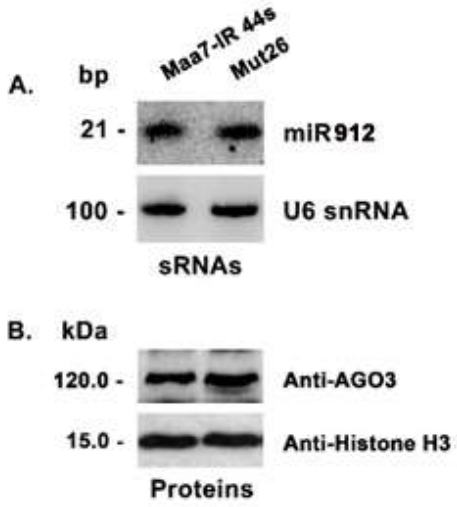

Figure 5.3. Analysis of miR912 expression and Ago3 protein abundance. (A) Northern blot analysis of Chlamydomonas mature miRNA in the indicated strains. The same membrane was reprobed with the U6 snRNA sequence as a control for equivalent loading of the lanes. (B) The Argonaute 3 (AGO3) abundance was examined by immunoblotting. Histone H3 was analyzed as a loading control.

3. *Mut26* is not defective in siRNA/miRNA biogenesis

An RNAi defective mutant in sRNA biogenesis failed to generate mature miRNAs and was showed to decrease the accumulation of Ago proteins (Voshall, Kim, Ma, Moriyama, & Cerutti, 2015; Yamasakia, Onishi, Kim, Cerutti, & Ohamaa, 2016). To confirm whether *Mut26* affects sRNA biogenesis pathways, we examined expression levels of endogenous miR912 and abundances of Ago3 proteins. However, *Mut26* had similar levels of miR912 and Ago3 proteins compared with the parental strain (Figure 5.3). These results implied that CCR4 proteins do not alter sRNA biogenesis pathways in Chlamydomonas.

4. *Mut26* and the poly(A) tail length of endogenous transcripts

In animals, miRNAs trigger both translational repression and degradation of target mRNAs. CCR4/NOT complexes are involved in transcript deadenylation, which promotes RNA decay (Fukao, Aoyama, & Fujiwara, 2015; Huntzinger & Izaurralde, 2011; Iwakawa & Tomari, 2015; Jonas & Izaurralde, 2015). To investigate whether the Chlamydomonas CCR4 homolog also affects mRNA deadenylation, the poly(A) tail length of two diagnostic transcripts was determined by the G/I tailing assay in three strains (CC-124, Maa7-IR 44s, and *Mut26*). The distribution of poly(A) tail lengths for the Maa7 and ACT1 (encoding action) transcripts were very similar between *Mut26* and its parental strain, Maa7-IR 44s (Figure 5.4). *Mut26* lacks a CCR4-like protein, which is part of the CCR4/NOT complex required for shortening of poly(A) tails, promoting subsequent decay of cytoplasmic mRNAs (Iwakawa & Tomari, 2015; Jonas & Izaurralde, 2015). However, *Mut26* showed only a very minor (if any) increase in the poly(A) tail

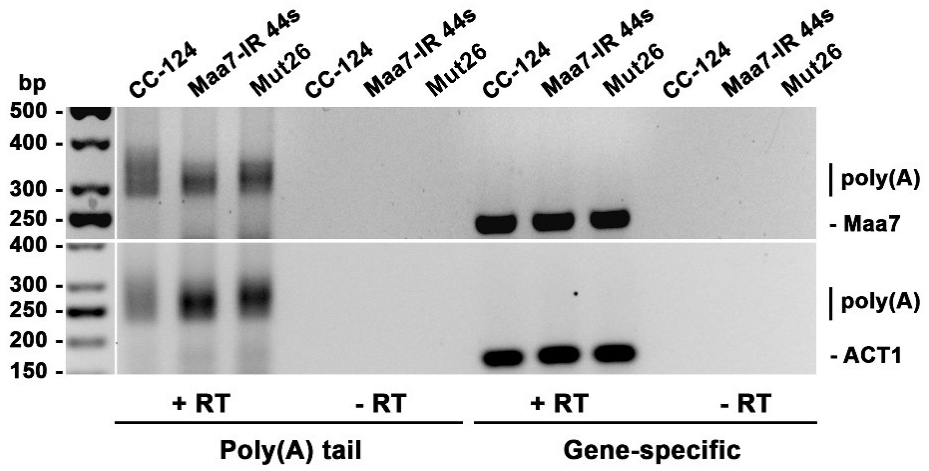

Figure 5.4. Analysis of polyadenylated tail lengths of the Maa7 and ACT1 transcripts of the indicated strains. Poly(A) tail length was examined using a G/I tailing protocol followed by RT-PCR. Reactions were performed as described in the methods in the presence (+RT) or absence (-RT) of reverse transcriptase.

length compared to Maa7-IR 44s in both siRNAs targeted transcripts (Maa7) and transcripts not regulated by RNAi (ACT1) (Figure 5.4). These results implied that CCR4 participates in sRNA-mediated translational repression in Chlamydomonas in a deadenylation-independent manner, consistent with its lack of effect on transcript decay of miRNA targets. Alternatively, a deadenylation defect caused by the deletion of CCR4 may be compensated by other components such as Pan2-Pan3, which is another cytoplasmic deadenylase.

4. Chlamydomonas CCR4 associates with NOT1, Caf1, eIF4A and DDX6 Proteins

To better understand the function of the CCR4 protein in Chlamydomonas, we identified interacting partners using a Co-immunoprecipitation approach (Figure 5.5). Affinity purification and mass spectrometry identification revealed several interacting

partners associated with *MUT26*, including NOT1 (the main body of the CCR4/NOT deadenylase), Caf1 (another subunit of the complex with ribonuclease activity), DDX6 (an RNA helicase and decapping cofactor), and eIF4A (a translational initiation factor) (Jonas & Izaurralde, 2015). We also identified as associated proteins eIF3A/C/M (eukaryotic translation Initiation factors-PCI domains) and several ribosomal proteins

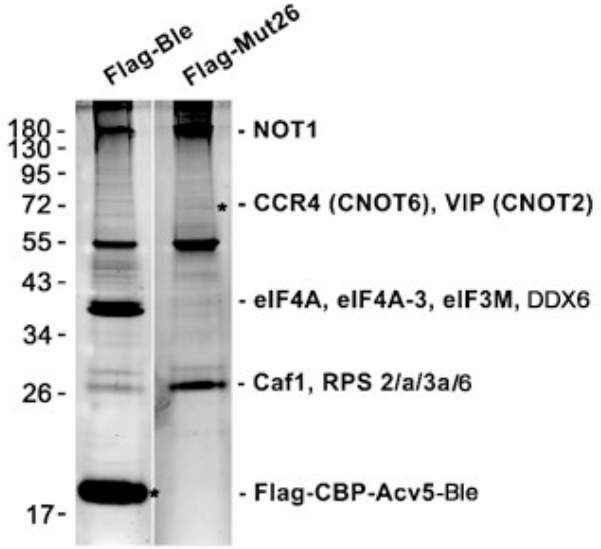

Figure 5.5. Co-immunoprecipitation of the Flag-tagged-MUT26 associated proteins. Epitope-tagged *Mut26*-associated proteins were pulled down and separated by 10% SDS-PAGE and visualized by Sypro-Ruby. Indicated interacting proteins were identified by mass spectrometry analyses using lysates of the tagged Ble (as a negative control) and *Mut26*-expressing strains. The asterisk mark indicates the Flag-tagged CCR4 protein.

(RPS 2/a/3/3a/5/6) (Table 5.1). These observations strongly suggested that Chlamydomonas CCR4 is indeed part of the canonical eukaryotic deadenylase complex and is also associated with translation initiation complexes and mRNA decay complexes. However, as discussed in chapter 1, the role of CCR4 in Ago-mediated translational repression is still unclear. To our knowledge, this is the first demonstration that CCR4 is

involved in sRNA-mediated translational repression in Viridiplantae. But, how CCR4 is recruited to the Ago complex and its mechanistic relationship remains to be explored.

Table 5.1. Proteins associated with the Flag-tagged CCR4 protein.

	Protein ID	M.W. (KDa)	Number of unique peptides		
			Exp. 1	Exp. 2	Exp. 3
CCR4/NOT complex	NOT1 (CNOT1)	155	22	0	33
	CCR4-like(CNOT6)	65	21	35	23
	Caf1 (CNOT8)	33	2	18	3
	VIP2 (CNOT2)	72	4	11	0
RNA helicases	DDX6	46	6	0	0
	eIF 4A (DDX2A)	47	0	17	4
	eIF 4A-3	46	0	5	0
PCI domain	eIF 3M	43	7	2	0
Ribosomal proteins	RPS2	30	3	0	0
	RPSa	31	3	0	0
	RPS3a	29	2	0	0
	RPS6	28	2	0	0

Discussion

sRNAs collectively control target mRNA stability and/or protein production by association with Ago proteins that guide them to their cognate mRNAs. sRNA/Ago effectors also recruit decapping and/or deadenylase complexes, triggering translational repression and, subsequently, mRNA degradation (Iwakawa & Tomari, 2015). Yet, the precise mechanisms for Ago-mediated translational repression remain unclear. For instance, the mode difference appears to depend on which cofactors are recruited by Ago,

whether GW182 is involved in the pathway, or the process depends on the organism and its developmental stages.

In animals, miRNA-mediated translational repression appears to occur in various ways (Nishimura *et al*., 2015). For example, miRNAs inhibit translation initiation by directly blocking the assembly of an eIF4F complex in Drosophila (Fukaya, Iwakawa, & Tomari, 2014). Biochemical assays showed that human miRNAs triggered the dissociation of eIF4AI/II from the target mRNA during translation initiation in a deadenylation-independent manner (Fukao *et al*., 2014). Ago proteins also recruit GW182 proteins that play essential roles in the translation inhibition and mRNA deadenylation steps (Iwakawa & Tomari, 2013). In this case, the Ago protein directly interacts with the GW182 protein and mRNA decay is immediately promoted by deadenylation via CCR4/NOT and Pan2/3 complexes (Braun, Huntzinger, Fauser, & Izaurralde, 2011; Jonas & Izaurralde, 2015). In plants, strikingly, GW182 orthologs do not appear to participate in the miRNA pathway (Iwakawa & Tomari, 2013; Li, 2013; Peter Brodersen, 2008; Yamasaki et al., 2013; Yang, Wu, & Poethig, 2012). In Arabidopsis, a plant-specific pipeline to identify proteins with GW/WG motifs failed to identify specific GW orthologs (Karlowski *et al*., 2010). These observations suggest that plant miRNA-mediated translational repression probably differs from the animal system or alternative mechanisms for the remittent of deadenylase complexes (GW182-independent) exist.

Our results provide direct evidence that CCR4 (and possibly the CCR4/NOT deadenylase complex) is indeed involved in sRNA-mediated translational repression in the unicellular green alga *Chlamydomonas reinhardtii*. It supports partial conservation of

the machinery involved in translational silencing between animal and plant systems. However, it is still unclear how sRNA-mediated translational repression occurs in the absence of GW182 and how Ago effectors may recruit the CCR4/NOT complex. sRNA-mediated translational repression in Chlamydomonas may provide a unique system for the characterization of the machinery involved in Viridiplantae and of its similarities and differences with that present in animals.

Materials and methods

1. Transgenic strains and culture conditions

In *Chlamydomonas reinhardtii*, Maa7 IR containing construct was transformed into a wild-type (CC-124) as previously described (Kim & Cerutti, 2009; Rohr, Sarkar, Balenger, Jeong, & Cerutti, 2004). Random insertional mutagenesis generated a RNAi-defective *Mut26* via a plasmid conferring resistance to paromomycin (Kindle, 1990) and the deletion of a CCR4 gene was confirmed by Tail-PCR (Singer & Burke, 2003). The full length of CCR4 coding sequences tagged with N-terminal FLAG-CBP-AcV5 were transformed into *Mut26*. The complementary strains, Mut26(MUT26)-07 and Mut26(MUT26)-42, were screened by Bleomycin antibiotic. Unless otherwise noted, all cells were grown in Tris-Acetate-Phosphate (TAP) medium under moderate light conditions (Harris, 1989).

2. RNA analyses

Total RNAs were isolated by TRI reagent (Molecular Research Center), following the manufacturer's instructions. For Northern blot, the purified RNAs were separated by agarose/formaldehyde gel electrophoresis and blotted onto Hybond-NX (GE Healthcare) nylon membranes that were hybridized with 32P-labeled probes (Rohr *et al*., 2004; Sambrook & Russell, 2001). For quantitative PCR analyses, DNase I-treated RNA samples were used as a template for first-strand cDNA synthesis, using an oligo(dT)18 primer and SuperScript III reverse transcriptase (Life Technologies). Primer pairs for the quantitative PCR amplifications were as follows: for *Cre17.g697550*, g697550-F (5'-GAGAGGATCGCGGACAACC-3') and g697550-R (5'-AGGACCGGTAGATGCTCTTGG-3'); for *Cre16.g683650*, g683650-F (5'-CAGTTTGAGCCCGACCTACG-3') and g683650-R (5'-CCACGCCGCACTCCAGC-3'); for *Cre01.g008950*, g008950-F (5'-GCTCAAAGGCAAGGGAAAAAT-3') and g008950-R (5'-TCCTTGAACATGCCCTCAATG-3'); for *Cre12.g552950*, g552950-F (5′-AACTGGATAGGCTGAGCAGGA-3′) and 552950-R (5'- TTGTGGGGACAGCTTCTTCTT-3'); and for Actin, ACT-Cod-F (5'-GACATCCGCAAGGACCTCTAC-3') and ACT Cod-R (5'GATCCACATTTGCTGGAAGGT-3'). The poly(A) tail length of specific transcripts was estimated using a G/I tailing protocol followed by RT-PCR analysis (Kusov, Shatirishvili, Dzagurov, & Gauss-Müller, 2001), according to a commercially available kit (USB; Affymetrix). The primer sequences for the poly(A) tail analyses were as follows: for ACT1, ACT-3'UTR-PF4 (5'-AAGATATGAGGAGCGGGTCA-3') and ACT-3'UTR-PR2 (5'-AAATGGTCCGAGCAGGTTTT-3'); and for Maa7, Maa7-3'UTR-PF1 (5'-GTGATTGAAAGGGGAGCGTA-3') and Maa7-3'UTR-PR1 (5'-ACATGCGATTGGTAGCAACA-3').

3. Protein analyses

For in vivo analysis, cells were grown to logarithmic phase and ~ 5×10^6 cells were pelleted by centrifugation and resuspended in 50 μL of SDS-gel running buffer. Then 10 μL aliquots were boiled and the samples were separated by 10-12% SDS-PAGE and electroblotted onto nitrocellulose membranes (Ma *et al*., 2013). The AcV5-tagged proteins were immunodetected by overnight incubation at 4°C with a 1:10,000 dilution of a mouse raised anti-AcV5 antibody (AbCam, Cambridge, MA, USA). After washing, the membrane was incubated with a rabbit anti-mouse secondary antibody, conjugated to horseradish peroxidase (HRP), for 2 hours at room temperature with a dilution of 1:10,000. For an Argonaute3 (AGO3) protein, electroblotted proteins were incubated at 4°C overnight with a 1:10,000 dilution of a rabbit antibody raised against the C-terminal peptide of AGO3 conjugated to KLH (CASRSGRGAGAAEGG; GenScript). The secondary antibody used was a goat anti-rabbit secondary antibody conjugated to HRP. For MAA7 protein, tryptophan synthase β (TSβ), an anti-TSβ antibody raised in a polyclonal goat (kindly provided by Thomas McKnight) at a 1:10,000 dilution was incubated at 4°C overnight. An anti-rabbit secondary antibody conjugated to horseradish peroxidase was used for autoradiography. The Chlamydomonas *Cre16.g683650* protein was immunodetected, following a standard procedure (Voshall et al., 2015), by overnight incubation at 4°C with a 1:10,000 dilution of a rabbit antibody raised against a C-terminal peptide (GIKPSAHKRGGVRM) conjugated to KLH (GenScript). A modification-insensitive polyclonal antibody (Abcam ab1791) was used to detect histone H3 used to adjust sample loading. A chemiluminescent substrate (Pierce, Rockford, IL, USA) was used for autoradiographic detection.

4. Co-immunoprecipitation

For the affinity purification experiment, 0.5~1.0 × 10^{10} cells were resuspended in a lysis buffer [20 mM Tris-HCl (pH 7.5), 150 mM KCl, 0.1 mM EDTA, 10% glycerol] containing 2.0 mM Benzamidine, 0.2 mM PMSF, and 5µl/ml plant protease inhibitor (Sigma). Cells were broken by two passages through a French-press at 3000 psi. An aliquot of the total lysate was stored (at -70°C) for SDS-PAGE and immunoblot analyses. The lysate was centrifuged at 16,000 g for 30 minutes at 4°C and the supernatant (Fraction S16) was centrifuged again at 100,000 g for 90 minutes at 4°C to separate out soluble proteins (Fraction S100). The supernatant was incubated with ANTI-FLAG (M2) beads (Sigma A2220) at 4°C for overnight and the beads were collected by centrifugation (500 g for 30sec). The beads were then washed with the following buffer for 15 minutes (each wash) at 4°C: Three washes with Washing Buffer B [20 mM Tris-HCl (pH 7.5), 150 mM KCl, 10% glycerol] containing 0.2 mM PMSF, 5µl/ml plant protease inhibitor (Sigma), and 0.1% Triton X-100. The beads were resuspended in 2X SDS sample loading buffer and the proteins in the supernatant were resolved on a 10% SDS-PAGE gel to separate the bound proteins with FLAG-AcV5-tagged CCR4 or AGO3 and AcV5-Flag-Ble detection as a control for a non-specific association of proteins with the beads. Subsequent mass spectrometry analysis was performed.

References

Ameres, S. L., & Zamore, P. D. (2013). Diversifying microRNA sequence and function. *Nat Rev Mol Cell Biol, 14*(8), 475-488. doi:10.1038/nrm3611

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell, 136*(2), 215-233. doi:10.1016/j.cell.2009.01.002

Braun, J. E., Huntzinger, E., Fauser, M., & Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell, 44*(1), 120-133. doi:10.1016/j.molcel.2011.09.007

Brodersen, P., Sakvarelidze-Achard, L., Bruun-Rasmussen, M., Dunoyer, P., Yamamoto, Y. Y., Sieburth, L., & Voinnet, O. (2008). Widespread Translational Inhibition by Plant miRNAs and siRNAs. *Science, 320*.

Chen, X. (2004). A MicroRNA as a Translational Repressor of APETALA2 in Arabidopsis Flower Development.

Duanmua, D., Millerb, A. R., Horkenb, K. M., Weeksb, D. P., & Spaldinga, M. H. (2009). Knockdown of limiting-CO2–induced gene HLA3 decreases HCO3 transport and photosynthetic Ci affinity in Chlamydomonas reinhardtii. *PNAS, 106*(14).

Fukao, A., Aoyama, T., & Fujiwara, T. (2015). The molecular mechanism of translational control via the communication between the microRNA pathway and RNA-binding proteins. *RNA Biol, 12*(9), 922-926. doi:10.1080/15476286.2015.1073436

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., . . . Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. *Mol Cell, 56*(1), 79-89. doi:10.1016/j.molcel.2014.09.005

Fukaya, T., Iwakawa, H. O., & Tomari, Y. (2014). MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. *Mol Cell, 56*(1), 67-78. doi:10.1016/j.molcel.2014.09.004

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

Huntzinger, E., & Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet, 12*(2), 99-110. doi:10.1038/nrg2936

Ibrahim, F., Rymarquis, L. A., Kim, E. J., Becker, J., Balassa, E., Green, P. J., & Cerutti, H. (2010). Uridylation of mature miRNAs and siRNAs by the MUT68 nucleotidyltransferase promotes their degradation in Chlamydomonas. *Proc Natl Acad Sci U S A, 107*(8), 3906-3911. doi:10.1073/pnas.0912632107

Iwakawa, H. O., & Tomari, Y. (2013). Molecular insights into microRNA-mediated translational repression in plants. *Mol Cell, 52*(4), 591-601. doi:10.1016/j.molcel.2013.10.033

Iwakawa, H. O., & Tomari, Y. (2015). The Functions of MicroRNAs: mRNA Decay and Translational Repression. *Trends Cell Biol, 25*(11), 651-665. doi:10.1016/j.tcb.2015.07.011

Jonas, S., & Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. *Nat Rev Genet, 16*(7), 421-433. doi:10.1038/nrg3965

Karlowski, W. M., Zielezinski, A., Carrere, J., Pontier, D., Lagrange, T., & Cooke, R. (2010). Genome-wide computational identification of WG/GW Argonaute-binding proteins in Arabidopsis. *Nucleic Acids Res, 38*(13), 4231-4245. doi:10.1093/nar/gkq162

Kim, E.-J., & Cerutti, H. (2009). Targeted Gene Silencing by RNA Interference in Chlamydomonas. *93*, 99-110. doi:10.1016/s0091-679x(08)93005-3

Kindle, K. L. (1990). High-frequency nuclear transformation of Chlamydomonas reinhardtii. *Proc. Natl. Acad. Vol. 87, pp. 1228-1232.*

Kusov, Y. Y., Shatirishvili, G., Dzagurov, G., & Gauss-Müller, V. (2001). A new G-tailing method for the determination of the poly(A) tail length applied to hepatitis A virus RNA. *Nucleic Acids Res., 29: E57–E7.*

Li, S. L., L. Zhuang, X. Yu, Y. Liu, X. Cui, X. Ji, L. Pan, Z. Cao, X. Mo, B. Zhang, F. Raikhel, N. Jiang, L. Chen, X. (2013). MicroRNAs inhibit the translation of target mRNAs on the endoplasmic reticulum in Arabidopsis. *Cell, 153*(3), 562-574. doi:10.1016/j.cell.2013.04.005

Ma, X., Kim, E. J., Kook, I., F. Voshall, A. Moriyama, & E. Cerutti, H. (2013). Small Interfering RNA-Mediated Translation Repression Alters Ribosome Sensitivity to Inhibition by Cycloheximide in Chlamydomonas reinhardtii. *Plant Cell, 25*(3), 985-998. doi:10.1105/tpc.113.109256

Miyoshi, T., Ito, K., Murakami, R., & Uchiumi, T. (2016). Structural basis for the recognition of guide RNA and target DNA heteroduplex by Argonaute. *Nat Commun, 7*, 11846. doi:10.1038/ncomms11846

Nishimura, T., & Fabian, M. R. (2016). Scanning for a unified model for translational repression by microRNAs. *EMBO J, 35*(11), 1158-1159. doi:10.15252/embj.201694324

Nishimura, T., Padamsi, Z., Fakim, H., Milette, S., Dunham, W. H., Gingras, A. C., & Fabian, M. R. (2015). The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs. *Cell Rep, 11*(9), 1425-1436. doi:10.1016/j.celrep.2015.04.065

Oey, M., Ross, I. L., Stephens, E., Steinbeck, J., Wolf, J., Radzun, K. A., . . . Hankamer, B. (2013). RNAi knock-down of LHCBM1, 2 and 3 increases photosynthetic H2 production efficiency of the green alga Chlamydomonas reinhardtii. *PLoS One, 8*(4), e61375. doi:10.1371/journal.pone.0061375

Peter Brodersen, L. S.-A., 1 Marianne Bruun-Rasmussen,1Patrice Dunoyer,1 Yoshiharu Y. Yamamoto,2 Leslie Sieburth,3 Olivier Voinnet1*. (2008). Widespread Translational Inhibition by Plant miRNAs and siRNAs.

Pollock, S. V., Prout-Jr, D. L., Godfrey, A. C., Lemaire, S. D., & Moroney, J. V. (2004). The Chlamydomonas reinhardtii proteins Ccp1 and Ccp2 are required for long-term growth, but are not necessary for efficient photosynthesis, in a low-CO2 environment. *Plant Molecular Biology, 56*(125–132).

Rohr, J., Sarkar, N., Balenger, S., Jeong, B. R., & Cerutti, H. (2004). Tandem inverted repeat system for selection of effective transgenic RNAi strains in Chlamydomonas. *Plant J, 40*(4), 611-621. doi:10.1111/j.1365-313X.2004.02227.x

Sambrook, J., & Russell, D. W. (2001). Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press, New York*.

Singer, T., & Burke, E. (2003). High-Throughput TAIL-PCR as a Tool to Identify DNA Flanking Insertions. In E. Grotewold (Ed.), *Plant Functional Genomics* (pp. 241-271). Totowa, NJ: Humana Press.

Song, J. J., Liu, J., Tolia, N. H., Schneiderman, J., Smith, S. K., Martienssen, R. A., . . . Joshua-Tor, L. (2003). The crystal structure of the Argonaute2 PAZ domain

reveals an RNA binding motif in RNAi effector complexes. *Nat Struct Biol, 10*(12), 1026-1032. doi:10.1038/nsb1016

Song, J. J., Smith, S. K., Hannon, G. J., & Joshua-Tor, L. (2004). Crystal Structure of Argonaute and Its Implications for RISC Slicer Activity.

Voshall, A., Kim, E. J., Ma, X., Moriyama, E. N., & Cerutti, H. (2015). Identification of AGO3-associated miRNAs and computational prediction of their targets in the green alga Chlamydomonas reinhardtii. *Genetics, 200*(1), 105-121. doi:10.1534/genetics.115.174797

Wilczynska, A., & Bushell, M. (2015). The complexity of miRNA-mediated repression. *Cell Death Differ, 22*(1), 22-33. doi:10.1038/cdd.2014.112

Yamasaki, T., Voshall, A., Kim, E. J., Moriyama, E., Cerutti, H., & Ohama, T. (2013). Complementarity to an miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga Chlamydomonas reinhardtii. *Plant J, 76*(6), 1045-1056. doi:10.1111/tpj.12354

Yang, L., Wu, G., & Poethig, R. S. (2012). Mutations in the GW-repeat protein SUO reveal a developmental function for microRNA-mediated translational repression in Arabidopsis. *Proc Natl Acad Sci U S A, 109*(1), 315-320. doi:10.1073/pnas.1114673109

Yuan, Y. R., Pei, Y., Ma, J. B., Kuryavyi, V., Zhadina, M., Meister, G., . . . Patel, D. J. (2005). Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage. *Mol Cell, 19*(3), 405-419. doi:10.1016/j.molcel.2005.07.011

Zamore, P. D., & Haley, B. (2005). Ribo-gnome: the big world of small RNAs. *Science, 309*(5740), 1519-1524. doi:10.1126/science.1111444

CHAPTER 6

CONCLUSION

Two major noncoding-small RNAs (21-24nt), miRNAs and siRNAs, have been shown to be essential regulators for gene silencing in eukaryotes (David P. Bartel, 2004). They only differ in the structure of precursors (a short hairpin or long double-stranded RNAs, respectively) and the origin of the genes (either endogenous and/or exogenous RNAs). Both sRNAs share many of these characteristics, including biogenesis and the cellular functions involved in gene silencing. Also, both sRNAs have a 2-nucleotide overhang at the 3'end when processed by Dicer and both prefer a uracil nucleotide at the 5'end bound to on Ago effector. Based on the above knowledge, these chapters investigated the mechanisms and functions of sRNA-dependent gene silencing in *Chlamydomonas reinhardtii*.

Gene silencing in microalgae, as reviewed in Chapter 2, described the core components of this process in Archaeplastida microalgae. We initially collected interesting protein sequences from almost finished or draft genome projects, including 14 species of microalgae. Phylogenetic trees showed that the key proteins of RNAi and DNA/histone methylation in gene silencing were evolutionally conserved in most microalgae. However, some domains were missing when compared to a higher plant, suggesting the precise mechanisms of gene silencing may be diverse. Chapter 2 provided useful knowledge for the development of the genetic engineering field and improved

feedstocks for biofuel production derived from microalgae, although many unsolved questions remain to understand RNAi machinery in Chlamydomonas.

Chlamydomonas miRNAs were first characterized by two research groups in 2007 (Molnar, Schwach, Studholme, Thuenemann, & Baulcombe, 2007; Zhao *et al*., 2007). To date, there are 50 families of Chlamydomonas miRNAs reported in the miRbase website. Most of Chlamydomonas miRNAs were not conserved in between species or among other kingdoms, suggesting a recent evolutionary origin (Moran, Agron, Praher, & Technau, 2017; Taylor, Tarver, Hiscock, & Donoghue, 2014). To address possible functions of Chlamydomonas miRNAs in cellular processes, we examined the level of miRNA expressions under different starvation conditions, as shown in Chapter 3. We found differentially expressed miRNAs in responses to nutritional changes, but no functional impacts of miRNAs.

Our laboratory developed a tandem inverted repeat (TIR) RNAi system for highly efficient integration of TIR-RNAi transgenes into the Chlamydomonas genome (as reviewed in Chapter 4). The TIR-RNAi system allows for the direct selection of effective transgenic RNAi lines and the simultaneous silencing of two unrelated genes (Figure 4.1). This RNAi system can be applied to analyze gene functions or manipulate metabolite production in *Chlamydomonas reinhardtii*.

Ago-bound sRNAs repress gene expression via translational repression. To address this issue, many researchers have investigated the translation initiation step by searching the interacting components of the Ago complex, such as the translation initiation complex eIF4F (Braun *et al*., 2011; Fukaya *et al*., 2014). After first inhibition during translational initiation, target mRNAs are directed toward the RNA degradation

process. However, the mechanisms of the intermediate step before complete degradation remain unknown. Lastly, little is known about the role of CCR4 in the translation repression mechanism and the GW182-independent mechanism. To approach these complicated mechanisms, we characterized the deadenylase CCR4 protein, which is part of the CCR4-NOT complex, in Chapter 5.

References

Bartel, D. P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. *Cell, 116*, 281-297.

Braun, J. E., Huntzinger, E., Fauser, M., & Izaurralde, E. (2011). GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell, 44*(1), 120-133. doi:10.1016/j.molcel.2011.09.007

Fukaya, T., Iwakawa, H. O., & Tomari, Y. (2014). MicroRNAs block assembly of eIF4F translation initiation complex in Drosophila. *Mol Cell, 56*(1), 67-78. doi:10.1016/j.molcel.2014.09.004

Molnar, A., Schwach, F., Studholme, D. J., Thuenemann, E. C., & Baulcombe, D. C. (2007). miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. *Nature, 447*(7148), 1126-1129. doi:10.1038/nature05903

Zhao, T., Li, G., Mi, S., Li, S., Hannon, G. J., Wang, X. J., & Qi, Y. (2007). A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii. *Genes Dev, 21*(10), 1190-1203. doi:10.1101/gad.1543507