Cooperative processing of primary miRNAs by DUS16 and DCL3 in the unicellular green alga *Chlamydomonas reinhardtii*

Tomohito Yamasaki and Heriberto Cerutti

*Division of Environmental Photobiology, National Institute for Basic Biology, Okazaki, Aichi, Japan; School of Biological Science and Center for Plant Science Innovation, University of Nebraska–Lincoln, Lincoln, NE, USA*

**ABSTRACT**

We have previously reported that the RNA-binding protein Dull slicer 16 (DUS16) plays a key role in the processing of primary miRNAs (pri-miRNAs) in the unicellular green alga *Chlamydomonas reinhardtii*. In the present report, we elaborate on the interaction of DUS16 with Dicer-like 3 (DCL3) during pri-miRNA processing. Comprehensive analyses of small RNA libraries derived from mutant and wild-type algal strains allowed the de novo prediction of 35 pri-miRNA genes, including 9 previously unknown ones. The pri-miRNAs dependent on DUS16 for processing largely overlapped with those dependent on DCL3. Our findings suggest that DUS16 and DCL3 work cooperatively, presumably as components of a microprocessor complex, in the processing of the majority of pri-miRNAs in *C. reinhardtii*.

**KEYWORDS**

Argonaute; *Chlamydomonas reinhardtii*; Dicer; miRNA; RNA-binding protein; small RNA-seq

MicroRNAs (miRNAs) are loaded into Argonaute (AGO) proteins during the formation of the RNA-induced silencing complex (RISC). The main function of miRNAs in RNA silencing is guiding RISC to target transcripts for inducing endonucleolytic RNA cleavage and/or translational repression. In general, miRNAs are embedded in long primary miRNA (pri-miRNA) transcripts containing stem-loop structures and have to be processed to mature miRNAs with the assistance of RNase III Dicer and associated RNA-binding proteins.

We have recently reported that in the unicellular green alga *Chlamydomonas reinhardtii*, an RNA-binding protein, Dull slicer 16 (DUS16), is required for pri-miRNA processing and associates with Dicer-like 3 (DCL3), which in turn is involved in the biogenesis of the majority of miRNAs. We also reported that AGO3, which is one of the 3 AGOs encoded in the *C. reinhardtii* genome, predominantly binds to mature miRNAs and determines miRNA-mediated post-transcriptional gene silencing. The present report contains a comprehensive analysis of our previously published small RNA-seq (sRNA-seq) data [from the *AGO3* mutant (ago3-1); the *DUS16* mutant (dus16-1); the parental strain of these mutants Gluc(1×), which expresses a reporter luciferase transgene in the wild-type background; and the wild-type strain CC-124] to predict de novo pri-miRNAs and gain insight into the functional coupling between DUS16 and DCL3.

From the sRNA-seq raw data of CC-124, Gluc(1×), ago3-1, and dus16-1, adaptor sequences were removed and reads ranging from 17 to 25 nucleotides in length were selected for further analyses. The alignment of sorted sRNA reads from the Gluc(1×) sRNA library to the *C. reinhardtii* genome (Ch_genome_v5.0) using miRA, an miRNA discovery tool for plants and algae, led to the identification of 1,062 inverted repeat loci encoding stem-loop RNAs. To stringently screen for genuine pri-miRNA genes, sRNA sequences with <10 read counts were excluded from the libraries, and the remaining redundant sRNA reads were aligned with *C. reinhardtii* gene models encompassing the inverted repeats using CLC genomic workbench. Gene models with <90 mapped-sRNA read counts in the sRNA libraries of CC-124 and Gluc(1×) and/or those without a predominant sRNA species on an arm of the predicted stem-loop structure were discarded.
Based on the above workflow, 35 gene models were annotated as pri-miRNA genes, including 9 previously unknown ones (Table 1).

A comparison of total sRNA read counts, mapped on the predicted pri-miRNA genes, from dus16–1 and Gluc (1×) revealed that the production of mature sRNAs from 33 of the 35 pri-miRNAs is significantly lower in dus16–1, suggesting that these pri-miRNAs are mainly processed in a DUS16-dependent manner (Table 1, Fig. S1). Twenty-four of the 35 identified miRNA genes were previously annotated as pri-miRNAs by Valli et al. and are predominantly processed by DCL3 (annotated as “high confidence,” “medium confidence” and/or “upregulated” in Table 1, Fig. S1).5 Furthermore, 22 of these 24 pri-miRNAs (91%) appear to require DUS16 for processing (Table 1; Fig. S1). This result suggests that, in addition to our previous finding of DUS16 physically interacting with DCL3,4 DUS16 is functionally coupled to DCL3, presumably as part of a microprocessor complex involved in the processing of the majority of C. reinhardtii pri-miRNAs.

On the other hand, 2 pri-miRNA transcripts corresponding to Cre04.g217925 and Cre06.g274550, which give rise to mature miR-1144 and miR-1162, respectively, are processed in a DCL3-dependent and DUS16-independent manner (Table 1). In the ago3–1 mutant, the number of mature sRNAs generated from these pri-miRNAs is very low, indicating that most likely, they are authentic pri-miRNAs (Table 1, Fig. 2, Fig. S1). Some

---

**Figure 1.** Model for miRNA biogenesis and action in Chlamydomonas reinhardtii. Dull slicer 16 (DUS16) recognizes nascent pri-miRNA transcripts (A). Dicer-like 3 (DCL3) mediates processing of most pri-miRNAs to miRNA duplexes with assistance of DUS16 (B). Argonaute 3 (AGO3) incorporates most Chlamydomonas mature miRNAs, having a U as their 5′ nucleotide, and forms the RISC (C). AGO3-RISC recognizes target transcripts and induces slicing and/or translational repression (D).

---

**Figure 2.** Frequency (counts) of small RNA (sRNA) reads matching the inverted repeat regions of Cre10.g444300 (A) and Cre06.g274550 (B) in the AGO3 mutant (ago3–1), the DUS16 mutant (dus16–1), and their parental strain Gluc(1×). Schematic diagrams of gene structures, indicating predicted start and stop codons, are shown at the bottom of each panel. Inverted repeat regions are indicated in red. Gray bars represent the coverage of sRNA read counts on the corresponding sequences.
Table 1. De novo prediction of primary and mature miRNAs.

| Gene ID | precursor | Gluc(100) | Gluc(16-1) | dd16-1 | Location of stem-loop (strand) | Length (nt) | MIR gene | Predicted as miRNA with (DG3 mutant) | Mature miRNA sequences | Length (nt) | agoa-3/-Kl8c (1/2) |
|---------|-----------|-----------|-----------|--------|--------------------------------|-------------|---------|-----------------------------------|------------------------|-------------|------------------|
| Cre01.g011500 | RNP1, 26S proteasome regulatory subunit | 13,070 12,102 12,596 2,102 1,541 1,972 | 0.01 | chromosome_3:3125402.32,257212 | 290 | intron | MIR9006 | CCGTGTCACTGCTAGTCAGG | 21 | 2.10 |
| Cre01.g023393 | no putative conserved proteins/ | 2,030 1,944 1,997 117 61 89 | 0.04 | chromosome_3:1,37224,932,27512 | 178 | 3’UTR | medium confidence | TGGACACTAGAACCAAGACACA | 22 | 0.50 |
| Cre01.g038390 | no putative conserved proteins/ | 3,881 3,515 3,698 84 54 69 | 0.02 | chromosome_3:7,74493,774,51159 | 1,183 | 5’UTR-exon | upregulated | TACATGATCCACTTGGAGG | 21 | 0.27 |
| Cre01.g143342 | no putative conserved proteins/ | 2,353 2,190 2,272 93 40 67 | 0.03 | chromosome_3:2,91274,921,92630 | 149 | 3’UTR | Cluster14712 | TGGTGCTGCTGCGCTCCCTAC | 21 | 1.70 |
| Cre01.g159590 | protein kinase | 7,435 696 7,199 128 89 109 | 0.02 | chromosome_3:50,738,982,673,4002 | 121 | 3’UTR | Cluster16441 | TGGTACTGCCGTTAATGTAAT | 21 | 0.51 |
| Cre01.g205230 | no putative conserved proteins/ | 2,279 2,071 2,175 327 166 247 | 0.11 | chromosome_3:73,760,188,73,766,45 | 728 | exon-intron | upregulated | TACGGGTCGCTGCTGAGACCC | 22 | 0.20 |
| Cre02.g217925 | KECLH repeat domain | 2,154 2,159 2,157 4,616 3,668 4,142 | 1.92 | chromosome_4:4,355,731,4,330,46 | 276 | intron-exon | MIR1144 | medium confidence | TGGGCGCTGCGTGCAGCG | 21 | 0.71 |
| Cre02.g220661 | no putative conserved proteins/ | 3,966 3,279 3,588 253 116 185 | 0.05 | chromosome_4:6,230,070,22,055,86 | 1,545 | 3’UTR | model | TGGGGCGCTGCTGACGACG | 22 | 0.14 |
| Cre04.g225700 | mediator of RNA polymerase II transcription subunit | 55,579 52,177 53,878 2,535 2,010 2,273 | 0.04 | chromosome_4:5,410,059,3,100,78 | 183 | Cluster17620 | high confidence | TGGGCGGCGCTGTTAGCAG | 22 | 0.16 |
| Cre05.g238343 | no putative conserved proteins/ | 17,163 15,132 16,748 206 117 162 | 0.01 | chromosome_5:5,298,442,2,960,73 | 1,293 | 3’UTR-exon | Cluster16906 | TGGCCATCGCTTGGAA | 21 | 0.09 |
| Cre05.g239900 | no putative conserved proteins/ | 57,037 54,236 55,637 790 595 693 | 0.03 | chromosome_5:3,332,276,32,327,68 | 121 | exon | Cluster10996 | AGGCGTTAGAGTGTAATG | 22 | 1.32 |
| Cre05.g242180 | no putative conserved proteins/ | 6,071 5,467 5,769 0 0 0 | 0.00 | chromosome_5:18,103,743,18,141,34 | 147 | 3’UTR | MIR913 | Cluster1793 | TGCCTGGCGTCGTTGACAG | 21 | 0.23 |
| Cre05.g242031 | no putative conserved proteins/ | 6,514 5,970 6,242 14 0 7 | 0.00 | chromosome_5:1,790,47,1,920,88 | 261 | 3’UTR | MIR913 | Cluster10996 | TGCCTGGCGTCTAGTACG | 21 | 0.98 |
| Cre05.g247100 | GoH3 spore coat protein | 4,209 3,519 3,864 760 492 426 | 0.16 | chromosome_5:3,953,231,3,955,515 | 285 | intron-exon | MIR918/919 | Cluster18100 | TGCCTGGCGTCTAGTACG | 21 | 0.19 |
| Cre06.g246052 | WD40 repeat domain | 19,750 18,386 19,068 349 202 276 | 0.01 | chromosome_5:6,880,412,22,7159 | 208 | intron-exon | MIR1196 | Cluster1996 | TACGGCAGGAGTAGAT | 21 | 0.06 |
| Cre06.g274530 | protein kinase | 12,480 11,752 12,051 19,016 14,460 16,253 | 1.38 | chromosome_6:3,609,217,306,749 | 93 | Cluster19538 | medium confidence | TGGTATGAGTGGCTGCT | 22 | 0.08 |
| Cre06.g278206 | lipoxigenase | 71,631 65,283 68,438 2,312 1,338 1,825 | 0.03 | chromosome_6:6,301,312,401,1518 | 198 | 5’UTR-exon | MIR9007 | AAGACTGCTCAGCCATCG | 20 | 0.37 |
| Cre06.g295330 | no putative conserved proteins/ | 3,134 2,618 2,966 103 64 84 | 0.03 | chromosome_6:6,885,401,658,427 | 264 | 3’UTR | Cluster15016 | TACGGCAGGAGTAGAT | 21 | 0.25 |
| Cre07.g312630 | no putative conserved proteins/ | 5,931 1,873 1,904 96 58 77 | 0.04 | chromosome_6:7,753,97,781,14 | 517 | 3’UTR | Cluster22537 | TGCCTGGCTGCTTCCTAGG | 20 | 0.33 |
| Cre08.g35835 | no putative conserved proteins/ | 12,763 12,236 12,495 412 269 341 | 0.03 | chromosome_7:8,121,841,1,121,961 | 121 | 3’UTR | Cluster23547 | TACGGCAGGAGTAGAT | 22 | 0.16 |
| Cre10.g444100 | no putative conserved proteins/ | 87,015 81,861 84,438 2,858 1,766 2,312 | 0.03 | chromosome_10:3,398,982,340,0009 | 148 | 3’UTR | MIR9897 | Cluster2675 | TACGGCAGGAGTAGAT | 21 | 0.13 |
| Cre10.g452700 | no putative conserved proteins/ | 31,037 28,233 30,635 1,941 1,280 1,561 | 0.05 | chromosome_10:45,896,637,45,988,30 | 194 | intron-exon | Cluster2725 | TGCCTGGCGTCTAGTACG | 21 | 0.56 |

(Continued on next page)
| Gene ID | encoded proteins/domains | rep 1^a | rep 2^a | mean | rep 1^a | rep 2^a | mean | Position of stem-loop (strand) | Length (nt) | Location of stem-loop | MIR gene | Voshall et al^5 | Predicted as miRNA precursor with^b | Uregulated in the DCL3 mutant | Mature miRNA miRNA sequences | Length (nt) | ago3–1 Gluc(1×) |
|---------|-------------------------|---------|---------|------|---------|---------|------|-----------------------------|------------|---------------------|----------|----------------|-------------------------------|-----------------------------|-----------------------------|------------|-----------------|
| Cre10.g463400 | no putative conserved proteins/ domains | 32.962 | 34.490 | 37.962 | 2.437 | 1.645 | 2.061 | 0.06 chromosone_15:16199729, 16199916 | 88 intron | Cluster 52100 | medium confidence | upregulated | ATCTGCCTGCTGCTGAG | 21 0.23 |
| Cre11.g467630 | conserved hypothetical protein | 12.468 | 12.717 | 12.688 | 0.217 | 0.217 | 0.217 | 0.217 chromosone_11:16184675, 16184782 | 108 intron | Cluster 52100 | medium confidence | upregulated | AAAGCTGCTGCTGCTGAG | 21 0.63 |
| Cre12.g536301 | no putative conserved proteins/ domains | 28.770 | 26.661 | 21.717 | 2.72 | 2.72 | 2.72 | 0.01 chromosone_12:6166877, 6167231 | 20 0.61 chromosome 12:6166877, 6167231 immigration | Cluster 52100 | medium confidence | upregulated | TGCCAGAGAGAGGGCGAC | 21 0.29 |
| Cre13.g576070 | conserved hypothetical protein | 17.130 | 15.285 | 16.208 | 1.52 | 1.95 | 1.24 | 0.01 chromosone_13:20001062, 20001077 | 146 3'-UTR | Cluster 52100 | medium confidence | upregulated | AAAGCTGAGTGGAGAAG | 20 0.66 |
| Cre13.g579030 | anaphase promoting complex subunit 1 | 10.506 | 10.219 | 10.640 | 0.332 | 0.233 | 0.233 | 0.03 chromosone_13:2301400, 2301727 | 328 3'-UTR | Cluster 52100 | medium confidence | upregulated | TGACTCTCACTCTACGCGC | 21 0.23 |
| Cre14.g518390 | translation elongation factor 3 | 40.709 | 36.425 | 38.567 | 1.913 | 1.313 | 1.625 | 0.04 chromosone_14:1191293, 1192047 | 755 intron | Cluster 52100 | high confidence | upregulated | ATGGAAGGTGGGCTACGCGC | 21 0.25 |
| Cre15.g668203 | no putative conserved proteins/ domains | 2.430 | 2.412 | 2.421 | 0.00 | 0.00 | 0.00 | 0.00 chromosone_15:16383378, 16383521 | 144 intron-exon | Cluster 52100 | medium confidence | upregulated | GGCGCGCTGACGCTGAG | 21 0.57 |
| Cre16.g668638 | no putative conserved proteins/ domains | 1.133 | 1.081 | 1.071 | 0.11 | 0.11 | 0.11 | 0.01 chromosone_16:31434870, 31435158 | 289 5'-UTR-exon-intron | Cluster 52100 | medium confidence | upregulated | TGCAGCTGACTGGTCATGG | 21 1.22 |
| Cre17.g667530 | no putative conserved proteins/ domains | 32.592 | 28.227 | 30.410 | 6.170 | 4.240 | 5.205 | 0.17 chromosone_17:194516, 194869 | 354 intron-exon-intron-exon | Cluster 52100 | medium confidence | upregulated | ATGCAGCCGCGCCGAG | 21 0.62 |
| Cre17.g667800 | chromosome segregation protein | 3.527 | 3.108 | 3.148 | 4.58 | 3.30 | 3.94 | 0.28 chromosone_17:228757, 228827 | 133 intron | Cluster 52100 | high confidence | upregulated | CGTCCTTCAATACTCAAA | 22 0.95 |
| Cre17.g715737 | FAP164, flagellar Associated Protein 164 | 60.726 | 54.207 | 57.497 | 28.321 | 23.151 | 24.836 | 0.43 chromosone_17:5152751, 5152951 | 201 5'-UTR | Cluster 52100 | medium confidence | upregulated | TGGGAGGCGGCGAAGTGTGAG | 22 0.41 |
| Cre17.g741601 | no putative conserved proteins/ domains | 12.965 | 11.544 | 12.325 | 0.69 | 0.69 | 0.69 | 0.02 chromosone_17:6144100, 6144226 | 127 5'-UTR | Cluster 52100 | medium confidence | upregulated | AGGCGCGCGGCGAG | 21 0.27 |
| Cre24.g756567 | conserved hypothetical protein | 3.094 | 2.846 | 2.970 | 2.76 | 2.11 | 2.44 | 0.08 scaffold_12:282169, 282227 | 159 3'-UTR | Cluster 52100 | medium confidence | upregulated | AGGCGAGCGCGGCGAAGCG | 22 0.42 |

Notes: ^aPhytozome (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii).
^bAbsolute sRNA read counts from the individual sRNA libraries that align to each gene model.
^cMean values of 2 replicates.
^dRatio of the means of abundant mature miRNAs in dus 6–1 over Gluc(1×).
^eLength of the sequences corresponding to a stem-loop RNA.
^fmiRBASE (http://www.mirbase.org/).
^gPreviously annotated pri-miRNA genes published by Voshall et al.^5
^hPreviously annotated pri-miRNA genes with high or medium confidence interval published by Valli et al.^5
^iPutative pri-miRNA genes with abundant upregulated transcripts in the DCL3 mutant (Valli et al).^5
^jRatio of the means of abundant mature miRNAs in ago3–1 over Gluc(1×).
sRNAs are also produced from the transcripts of inverted repeats in a DCL3-independent manner. These results imply the presence of minor DUS16- and/or DCL3-independent pri-miRNA-processing pathways in *C. reinhardtii*.

*C. reinhardtii* appears to possess canonical miRNA biogenesis pathways and miRNA-mediated post-transcriptional gene regulation with certain similarities to those in animals and plants. Mutant analyses revealed that the initial processing of the majority of pri-miRNAs relies on a putative microprocessor complex comprising both DUS16 and DCL3. In addition, our analyses also uncovered a minor set of pri-miRNAs that are likely processed in a DUS16 and/or DCL3-independent manner.

**Accession numbers**

Small RNA-seq raw data has been deposited in the DDBJ sequence read archive (DRA) under accession numbers DRA003930 and DRA004107 (CC-124 replicate #1, DRX040414; CC-124 replicate #2, CCCR040415; Glc1(×) replicate #1, DRX040416; Glc1(×) replicate #2, DRX040417; ago3−1 replicate#1, DRR045098; ago3−1 replicate#2 DRR045099; dus16−1 replicate #1, DRX043778; and dus16−1 replicate #2, DRX043779).

**Disclosure of potential conflicts of interest**

No potential conflicts of interest were disclosed.

**Acknowledgments**

We would like to thank the Functional Genomics Facility, NIBB Core Research Facilities for technical support.

**Funding**

This work was supported by NIBB Collaborative Research Program 15−103 (to T.Y.), JSPS Grant-in-Aid for Young Scientists (B) 16K18480 (to T.Y.), and a grant from the National Science Foundation (to H.C.).

**ORCID**

Tomohito Yamasaki http://orcid.org/0000-0001-9157-0209

**References**

[1] Meister G. Argonaute proteins: functional insights and emerging roles. Nat Rev Genet 2013; 14:447−59; PMID:23732335; http://dx.doi.org/10.1038/nrg3462

[2] Ha M, Kim VN. Regulation of microRNA biogenesis. Nat Rev Mol Cell Biol 2014; 15:509−24; PMID:25027649; http://dx.doi.org/10.1038/nrm3838

[3] Xie M, Zhang S, Yu B. microRNA biogenesis, degradation and activity in plants. Cell Mol Life Sci 2015; 72:87−99; PMID:25209320; http://dx.doi.org/10.1007/s00018-014-1728-7

[4] Yamasaki T, Onishi M, Kim EJ, Cerutti H, Ohama T. RNA-binding protein DUS16 plays an essential role in primary miRNA processing in the unicellular alga *Chlamydomonas reinhardtii*. Proc Natl Acad Sci USA 2016; 113:10720−5; PMID:27582463; http://dx.doi.org/10.1073/pnas.1523230113

[5] Valli AA, Santos BA, Hnatova S, Bassett AR, Molnar A, Chung BY, Baulcombe DC. 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 2016; 26:519−29; PMID:26968199; http://dx.doi.org/10.1101/gr.199703.115

[6] Yamasaki T, Kim EJ, Cerutti H, Ohama T. Argonaute3 is a key player in miRNA-mediated target cleavage and translational repression in *Chlamydomonas*. Plant J 2016; 85:258−68; PMID:26686836; http://dx.doi.org/10.1111/tpj.13107

[7] Evers M, Huttner M, Dueck A, Meister G, Engelmann JC. miRA: adaptable novel miRNA identification in plants using small RNA sequencing data. BMC Bioinformatics 2015; 16:370; PMID:26542525; http://dx.doi.org/10.1186/s12859-015-0798-3

[8] Molnár A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC. miRNAs control gene expression in the single-cell alga *Chlamydomonas reinhardtii*. Nature 2007; 447:1126−9; PMID:17538623; http://dx.doi.org/10.1038/nature05903

[9] Zhao T, Li G, Mi S, Li S, Hannon GJ, Wang XJ, Qi Y. A complex system of small RNAs in the unicellular green alga *Chlamydomonas reinhardtii*. Genes Dev 2007; 21:1190−203; PMID:17705533; http://dx.doi.org/10.1011/gad.1543507

[10] Cerutti H, Ibrahim F. Turnover of Mature miRNAs and siRNAs in Plants and Algae. Adv Exp Med Biol 2011; 700:124−39; PMID:21755478; http://dx.doi.org/10.1007/978-1-4419-7823-3_1

[11] Yamasaki T, Voshall A, Kim EJ, Moriyama E, Cerutti H, Ohama T. 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 2013; 76:1045−56; PMID:24127635; http://dx.doi.org/10.1111/tpj.12354

[12] Voshall A, Kim EJ, Ma X, Moriyama EN, Cerutti H. Identification of AGO3-associated miRNAs and computational prediction of their targets in the green alga *Chlamydomonas reinhardtii*. BMC Genomics 2012; 13:108; PMID:22439676; http://dx.doi.org/10.1186/1471-2164-13-108