Assembly of mitochondrial complex I requires the low-complexity protein AMC1 in *Chlamydomonas reinhardtii*

Nitya Subrahmanian[*†], Andrew David Castonguay[*‡], Claire Remacle[§], Patrice Paul Hamel[**].

*The Ohio State University, Department of Molecular Genetics, 500D Aronoff Laboratory, 318 W. 12th Avenue, Columbus, OH, 43210, USA.

† Plant Cellular and Molecular Biology Graduate Program, The Ohio State University, Columbus, Ohio, USA.

‡ Molecular Genetics Graduate Program, The Ohio State University, Columbus, Ohio, USA.

§University of Liège, Genetics and Physiology of Microalgae, UR InBios/Phytosystems, 4000 Liège, Belgium.

** The Ohio State University, Department of Molecular Genetics and Department of Biological Chemistry and Pharmacology, 500D Aronoff Laboratory, 318 W. 12th Avenue, Columbus, OH, 43210, USA.

SUPPORTING INFORMATION

Supporting Figures

Supporting Tables

Supporting Methods

SUPPORTING FIGURES

Figure S1. Sequence of the insertional *amc1-2* mutation in the *Cre16.g688900* gene of the *amc11* strain.

Black font depicts the section of exon 2 in the *Cre16.g688900* gene harboring the insertional *amc1-2* mutation in the *amc11* strain. **Orange bold font** indicates the insertional cassette conferring Hygromycin B resistance. The first 123 bp of the cassette, encoding the promoter region, was deleted upon insertion at this site. The **purple bold font** identifies the co-integrated herring sperm DNA sequences at the 5'- and 3'- ends of the cassette. The nucleotides with a black underline indicate the regions sequenced from TAIL-PCR or diagnostic PCR amplicons.

TGCATAGCTCCTGCCTTCTCACACAGTGGCTGGCGCGCAACGGCAGTAGCAACTG
GGACCTGTACCGGCTGCTACTGCGGCACCAGGCCGCGGCGGCGGCGGCGGCCG
CCGCAAGGCCCGATTTGGATACCATTGCAGTCGCGGTGCCGGGAACCGGCGGTC
AGGCTGGGATAAGTGACAATTTGAGAACATGTGCGGCGACGGCTGTGGCGAAAGC
GGTCTTGCAGCACAGCCGGAATGCAGCCCTCAAGCGCACCGGGCCCAACGGAAC
CGGTAGCGGTGACAGCGGATGGGTGGAGCCGCGGGCTGCGCTGCGACTGACGC
AGGCGCTGCGAGCGGCTCTGCCTTCCGCTACGGTTTCCCGCGGCAGCGGCGACG
CTTGGACCAGCCGCGGCGACGGCGCTGGTGCGGTGGTGGCGCCGGTGCCACGT
CAGGTGCGCGAGGCGGTGCGGGAGCTGCTGCTTGCTGCCGCCGCCCCCACCGG
GGCCCTCGGAGGCGGCAGCAGCGGTGCCGCAGGCGCTGCAGCAGGTGTTGGAG
CGGCGGGCGACGCAGTGGCGGCGGCCGGCCGGGAGGTGAGCTTCTGGATACGG
ACGGCGGCGGCCTTCACGTTGCCGCCGCACCCGGTCGGAGCGGCCGAGGCGGA
GTTGGTGGCGGGGCTGGGACCGGAGGCGGCGTTGGGTGGCGATCTGGCTGCGG
CCGCGGCGCCGCTGTTCCGCCGCGTCAGCCCGGACGCGCTGTTGGCCGCCGTCA
CCGACCTGGTACCTGGCGCGCAGCAACAGCAGCAACAGCAGCAACAGCAGCAAC
AGCAGCAACAGCAGCAACAGCAACAGCAGCAGCAGCAGGGCCCTGGCCT**AACAC**
GCACTGTAGCACACAAGTAAACGTTTGGCAAGCTTTTCAGGAATAGCGTGAACT
GAAAGAAAAAGAAGACTGCAAACCAGACGCCGACTCGGCCGTGTTGCTTTTGAA
ATTGAAAGTACCCTTGGGTGAACTTTGTCTTCGTGGTATTCTTGATTCTGATAGTC
TTTGTACAAATGTGTTTGCTCAACCATTTAGTTGTGAGCCCTGTAAAAAATGTTTT

CTCGACCGTTTTGTTGTGAGCCCTGTATGTTTCTAGCTTGCTTGCAACCATATAAA
CACCTTTGTTTCCATGGGTGTTTTGCTGTTCGTCTGTCCCCCAGATACAGACTGAA
TCTCCGAATCCAGGTTCTTGTTTATTTAGATTATTATGTTGGCCAAACCTCTTACA
CTGATTGTTCTGTTCAACCTAAAATAAAACAATTATAATCTAGGATATAAATTCTC
CCCGTCAACTATAAAAAAGGATGGATTTATGTTTATTGCAATACAAATACACCCTT
TTAAAAATGTATAACCTTGCCTACACTTTATGAACAACCCCTGTCTCCGTCCACCT
GCGACGTTGCATGGGCGCTCCGATGCCGCTCCAGGGCGAGCGCTGTTTAAATAG
CCAGGCCCCCGATTGCAAAGACATTATAGCGAGCTACCAAAGCCATATTCAAAC
ACCTAGATCACTACCACTTCTACACAGGCCACTCGAGCTTGTGATCGCACTCCGC
TAAGGGGGCGCCTCTTCCTCTTCGTTTCAGTCACAACCCGCAAACATGACACAAG
AATCCCTGTTACTTCTCGACCGTATTGATTCGGATGATTCCTACGCGAGCCTGCG
GAACGACCAGGAATTCTGGGAGGTGAGTCGACGAGCAAGCCCGGCGGATCAGG
CAGCGTGCTTGCAGATTTGACTTGCAACGCCCGCATTGTGTCGACGAAGGCTTTT
GGCTCCTCTGTCGCTGTCTCAAGCAGCATCTAACCCTGCGTCGCCGTTTCCATTT
GCAGCCGCTGGCCCGCCGAGCCCTGGAGGAGCTCGGGCTGCCGGTGCCGCCGG
TGCTGCGGGTGCCCGGCGAGAGCACCAACCCCGTACTGGTCGGCGAGCCCGGC
CCGGTGATCAAGCTGTTCGGCGAGCACTGGTGCGGTCCGGAGAGCCTCGCGTC
GGAGTCGGAGGCGTACGCGGTCCTGGCGGACGCCCCGGTGCCGGTGCCCCGCC
TCCTCGGCCGCGGCGAGCTGCGGCCCGGCACCGGAGCCTGGCCGTGGCCCTAC
CTGGTGATGAGCCGGATGACCGGCACCACCTGGCGGTCCGCGATGGACGGCAC
GACCGACCGGAACGCGCTGCTCGCCCTGGCCCGCGAACTCGGCCGGGTGCTCG
GCCGGCTGCACAGGGTGCCGCTGACCGGGAACACCGTGCTCACCCCCCATTCC
GAGGTCTTCCCGGAACTGCTGCGGGAACGCCGCGCGGCGACCGTCGAGGACCA
CCGCGGGTGGGGCTACCTCTCGCCCCGGCTGCTGGACCGCCTGGAGGACTGGC
TGCCGGACGTGGACACGCTGCTGGCCGGCCGCGAACCCCGGTTCGTCCACGGC
GACCTGCACGGGACCAACATCTTCGTGGACCTGGCCGCGACCGAGGTCACCGG
GATCGTCGACTTCACCGACGTCTATGCGGGAGACTCCCGCTACAGCCTGGTGCA
ACTGCATCTCAACGCCTTCCGGGGCGACCGCGAGATCCTGGCCGCGCTGCTCGA
CGGGGCGCAGTGGAAGCGGACCGAGGACTTCGCCCGCGAACTGCTCGCCTTCA
CCTTCCTGCACGACTTCGAGGTGTTCGAGGAGACCCCGCTGGATCTCTCCGGCTT
CACCGATCCGGAGGAACTGGCGCAGTTCCTCTGGGGGCCGCCGGACACCGCCC

CCGGCGCCTGATAAGGATCCCCGCTCCGTGTAAATGGAGGCGCTCGTTGATCTG
AGCCTTGCCCCCTGACGAACGGCGGTGGATGGAAGATACTGCTCTCAAGTGCTG
AAGCGGTAGCTTAGCTCCCCGTTTCGTGCTGATCAGTCTTTTTCAACACGTAAAA
AGCGGAGGAGTTTTGCAATTTTGTTGGTTGTAACGATCCTCCGTTGATTTTGGCCT
CTTTCTCCATGGGCGGGCTGGGCGTATTTGAAGCGGGTACCGGGCCCGTCATCC
CATGGAAGCTTCTTGCCGGGTGCACAGGCCTGTCACTTGTGCCTGTCTTAGTTGC
AGCAGTCGCCGGTGGACGGACCCTGGGCCTCGCCGGCCAGCACTGTGGACTGC
GCCGTGGGCGCCATGGATGCCCTGCTGCTGGACGCCGGCGCTGGCCAGCAACAC
CACCGGCAACAGCACCAACAGCACCCGCAACAGCACCAACAGCACCAAACTTCAC
ACCCGCAACAG

Figure S2. Sequence of the insertional *amc1-1* mutation in the *Cre16.g688900* gene of the *amc1* strain.

Black font depicts the section of exon 2 in the *Cre16.g688900* gene harboring the insertional *amc1-1* mutation in the *amc1* mutant. **Blue bold font** indicates the 1,536 bp of intergenic region from chromosome 17 inserted in exon 2. The **green bold font** and the **orange bold font** identify the co-integrated sequences at the 5'- and 3'- ends of the insertion site, respectively. The co-integrated sequence at the 5'-end (**green bold font**) has similarity to intron 24 of *Cre16.g695800*. The entire nucleotide sequence presented here was sequenced from TAIL-PCR and diagnostic PCR amplicons.

TGCATAGCTCCTGCCTTCTCACACAGTGGCTGGCGCGCAACGGCAGTAGCAACTG
GGACCTGTACCGGCTGCTACTGCGGCACCAGGCCGCGGCGGCGGCGGCGGCCG
CCGCAAGGCCCGATTTGGATACCATTGCAGTCGCGGTGCCGGGAACCGGCGGTC
AGGCTGGGATAAGTGACAATTTGAGAACATGTGCGGCGACGGCTGTGGCGAAAGC
GGTCTTGCAGCACAGCCGGAATGCAGCCCTCAAGCGCA**TAAACTCACTCTGACTC**
TCTCATGGCTGTCATCCGAGCCCCAGACTCTCCCCTCTCCCTATCTCTCCCCTAC
CCCACACCGCTGTGCCTGTTGCCCCCGACCTATCCCCGTAAGTCCTGAGCCGTTA
CCCGCAGCCCGTCAGTGATAGCCAAGGCCCCTGCCCACACTAGGGTCAAGAGTA
CCCCAACCCGCATAGCGTACGCCAGCAGAGCCTCTGCGCCAGGTTGCATACGTG
TAGTGTCTGGCAGCGATTAGGTACAGCAGCTGATAGCAGGGCGAGCTGGGAGC
GAGCCAGAGCAGGGTCCGGTTCGGTAGTTGGGCCTGAGCCCGTTCGGTTAGTGG
CAGTAGCAGGGAGTGAGGGAGACAGAGACGGGCGTAAGATAGCGTGAGATAGT
ACAGGGCCTGGGCCTAGGCAGTGGGGCGGACGGCATTGGACCGGGCACGCAC
GCCTTGTCTGCGGGTTGGGTCGTCGGGGGCTCTCCCCCCCCTTTCCCTTGGGGT
GTTCGCCTCCCCTGCCCCAGAGTGCAGGTCGACTGCGACACCTCTGGTCTGAAG
AAGTTTTGCTTATCCCTGCACCTGGTCTACCTGACCTGCGTGTGACCAAGTCCCC
GACACGCCGTGCCGGCTTGCGGGACCTGCTTGCTTTCCTGTCTCCCCTCTCTTCT
CCCTTCCCGTCTCTACCACCCGCCGCGTGCCCCGCACGCGAAGGGTAACCCCAG
CTACCGCCGGCGGCGGTGACTCTACTTGGGCAACCCGCCCGACCTTATCTCTCT
AATCTCTGATTACCCTGGGCGGCAAGGTTGAGTGAGACAGCCTTTGCCACGGCT
GTGCACCTGCACGCACGCCACCCATCCGGCTCTGTACGAGCCCCTGCTCCTGTT
ATCCAACCCGCCTAGACCAAACAGCGAAGCCCCTCCCGTGAGTAAAACAGCTAC

CGACGTCCCCATACGGCGGCGCTGGGCACTGCTGGGATCGAAGAGCCCCCTCT
ACGGCACGCGATCTGCCCCGACAGCTGTTGACACGCGGCGATAAAGTTAGGCG
AAAAAAGTGTCCTGTCATCTCTGGGGGGTCGGCACGCCTTTGTTGGTTTCGCTAC
CGGGACTTCTTCCATCCCCCAGTAGCCACTCATGTCAGACGAAGATGGCTTTCCC
CGCCGAGCGCGGCCTAGCGCTTTCCATACCGATCTGCAAGGACTTGACTATGAA
CGTATCACCCAACACTGCGAGCGCTTCGACCCGCAGTGGCATTCTGTGCGTGAT
TGACTGGAGTCCGTAGATGAGTTGGCAGAGGAATTCGGTTGGACTGAGTCACGA
AAGCTCCAACTTGCCTCTAAACGGCTTGGCCCCACGGCCAAGGAGTGGTACCGG
TCATGGAAGCTAGCCAACCCCATGGCATCGTCCACTTGGGAGGGCTTCTGTGAC
GCCCTGCGCGAGCGCTGGGGTGTGTCAGACCGCGAGCTGCATCTGGCCCTCGC
CAACTGCACACAAGGCCCCAAGGAGACGGTGCGCGAGTACGCCGACCGTTACC
TGGGCCTCGTGACCCAACTGCGCCTCGACTATAACCATCCGAGTGACCGGGCCC
AACGGAACCGGTAGCGGTGACAGCGGATGGGTGGAGCCGCGGGCTGCGCTGCG
ACTGACGCAGGCGCTGCGAGCGGCTCTGCCTTCCGCTACGGTTTCCCGCGGCAG
CGGCGACGCTTGGACCAGCCGCGGCGACGGCGCTGGTGCGGTGGTGGCGCCGG
TGCCACGTCAGGTGCGCGAGGCGGTGCGGGAGCTGCTGCTTGCTGCCGCCGCCC
CCACCGGGGCCCTCGGAGGCGGCAGCAGCGGTGCCGCAGGCGCTGCAGCAGGT
GTTGGAGCGGCGGGCGACGCAGTGGCGGCGGCCGGCCGGGAGGTGAGCTTCTG
GATACGGACGGCGGCGGCCTTCACGTTGCCGCCGCACCCGGTCGGAGCGGCCG
AGGCGGAGTTGGTGGCGGGGCTGGGACCGGAGGCGGCGTTGGGTGGCGATCTG
GCTGCGGCCGCGGCGCCGCTGTTCCGCCGCGTCAGCCCGGACGCGCTGTTGGC
CGCCGTCACCGACCTGGTACCTGGCGCGCAGCAACAGCAGCAACAGCAGCAACA
GCAGCAACAGCAGCAACAGCAGCAACAGCAACAGCAGCAGCAGCAGGGCCCTGG
CCTGGGTGAGCAACCGCGTGCCAGGAAGGGACCCTCGCTGTCAT

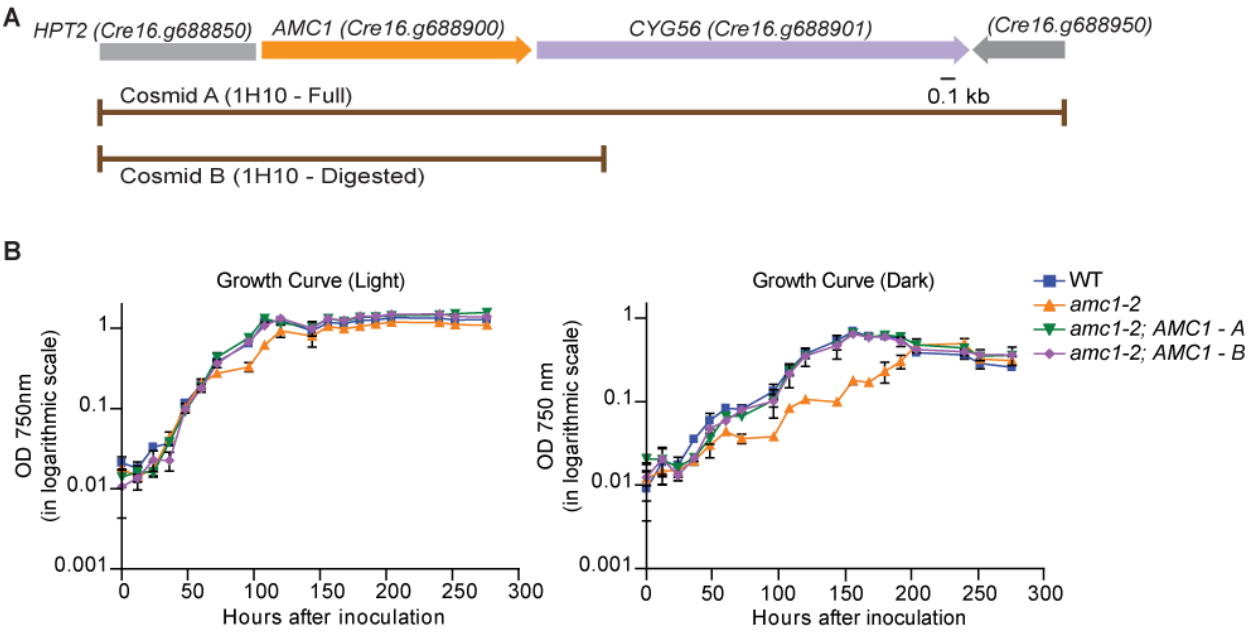

Figure S3. Rescue of the SID phenotype with cosmids containing the wild-type *AMC1* genomic DNA.

(A) Simplified map of the *Chlamydomonas* genomic DNA contained in the cosmids used for complementation studies. The 1H10 (Cosmid A) carries the truncated *HPT2* gene (*Cre16.g688850*), the *AMC1* gene (*Cre16.g688900*), the *CYG56* gene (*Cre16.g688901*), and *Cre16.g688950*. Cosmid 1H10 was digested by *Bam*HI restriction enzyme to obtain a truncated cosmid (Cosmid B) that only retains one full-length gene, namely *AMC1*. Both cosmids were used to transform the *amc1-2* mutant strain *amc11 (10G11)*.

(B) The growth of WT, *amc1-2* [*amc11 (10G11)*], [*amc1-2; AMC1-A*], and [*amc1-2; AMC1-B*] was documented by measuring optical density at A_{750}, in the light or in the dark, in the presence of acetate as a carbon source. The average of three biological replicates is reported here, with error bars indicating standard deviation of the mean. The generation time calculated from these growth curves is displayed in Figure 3A.

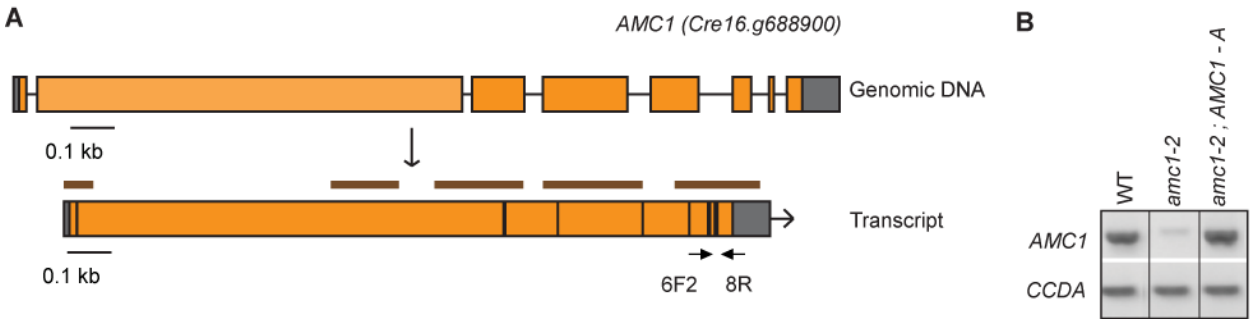

Figure S4. Transcript levels in the *amc1-2* mutant are restored by *AMC1* genomic DNA.

(A) The current *Chlamydomonas* genome database JGI v5.5 predicts the model for the *AMC1* gene as represented in the top panel. This gene model was experimentally corroborated by PCR amplification using *AMC1*-specific primers with total cellular cDNA prepared from the wild-type (4C$^-$). The brown lines, in the bottom panel, identify the sections of the *AMC1* cDNA successfully amplified and confirmed by sequencing. The following primer pairs were successful in amplifying the sequence of the *AMC1* cDNA: Cre16.g688900 5′UTR-F1/E2R10; 10G11 exon2F (1) / amc11-del1R; Cre16.g688900 exon2F4/exon3R2; au5.g6830 exon3F / exon4R; Cre16.g688900 exon4F2 / exon4-3R; Cre16.g688900 exon5-F3 / exon-8R; Cre16.g688900 exon8-1F / Cre16.g688900 3′UTR-1R; Cre16.g688900 exon8-1F / 3′UTR-2R (Table S2).

(B) Transcript abundance of *AMC1* was analyzed by RT-PCR in the wild-type (4C$^-$), *amc1-2* [*amc11 (10G11)*] strain, and the complemented strain [*amc1-2; AMC1-A*]. The *AMC1*-specific primers (au5.g6830 6F2 / Cre16.g688900 8R), used for amplification, are represented in (Figure S4A) with black arrows. *CCDA,* a gene involved in photosynthesis (GABILLY *et al.* 2011), was used as control for constitutive expression. Three independent biological replicates were performed, and one representative is shown in this figure. The vertical black lines indicate assembly of different lanes from the same gel.

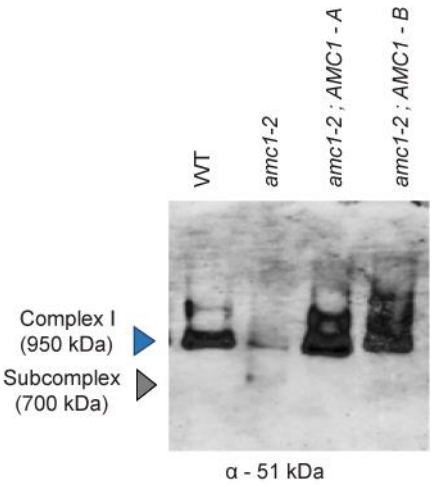

Figure S5. The *amc1-2* mutant display traces of a fully-assembled complex.
Immunoblotting using anti-51 kDa antibody of complexes separated by BN-PAGE. Two hundred µg of protein were loaded per lane. While the labile 700 kDa subcomplex can always be detected in the *amc1-2* mutant, in occasional membrane extractions, low levels of a fully-assembled complex I were also detected.

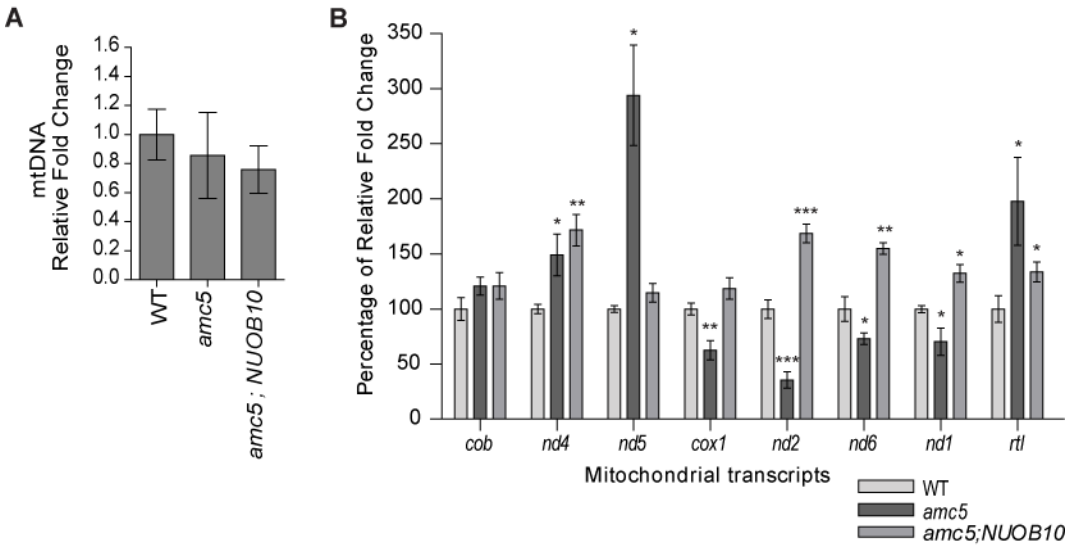

Figure S6. Mitochondrial transcript levels in a nuclear mutant characterized by the loss of the complex I subunit NUOB10.

(A) Real-time quantitative PCR (qPCR) was used to assess the relative quantity of mtDNA. The mitochondrial *nd4* gene was used as a target gene and the nuclear gene *TUA2,* encoding the alpha tubulin 2 protein, was used as the reference gene. The average was obtained from two biological replicates, each including three technical replicates, and the error bars represent standard deviation of the mean. The results are represented as fold change relative to WT (WT set to 1.0).

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

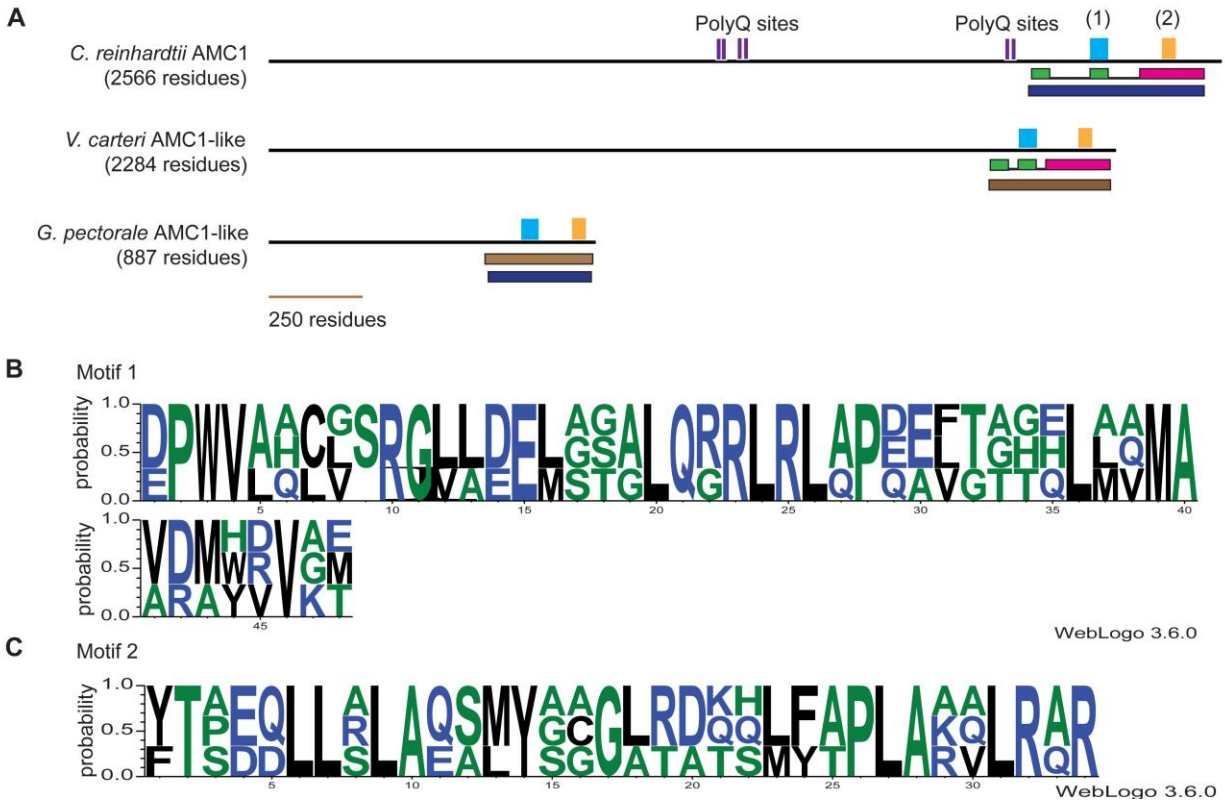

Figure S7. The C-terminus of AMC1 is similar to that of hypothetical algal proteins in *Volvox carteri* and *Gonium pectoral.*

Searches for AMC1-like proteins, via BLASTp (ALTSCHUL *et al.* 1990) and DELTA-BLAST (BORATYN *et al.* 2012), revealed that the AMC1 C-terminus (202 residues) has 37% identity with the C-terminus of another large hypothetical protein (2284 aa) from a closely related chlorophycean green alga *Volvox carteri* (Phytozome 12 Volvox database v2.1 ID: Vocar.0023s0109.1 / NCBI Accession: XP_002952084.1). AMC1 and the hypothetical *Volvox* protein show similarity at the C-terminus with another putative protein (887 aa in length) from the green alga *Gonium pectorale* (GPECTOR_23g121 / NCBI accession KXZ49034). While AMC1 shows only a weak similarity (29% identity amongst 432 residues), the *Volvox* hypothetical protein shows a significant similarity (36% identity amongst 339 residues) with the C-terminal region of the *Gonium* hypothetical protein.

(A) Conserved C-terminal motifs and polyQ repeats in the AMC1 protein sequence are shown. The polyQ repeats shown in Figure 7 are represented here as purple bars. Similarity searches with AMC1 identified similarity in the C-terminal end within a *Volvox*

carteri hypothetical protein. The pink box depicts 46% similarity in a 202 amino acid region, whereas the green boxes depict 64-68% in a 45 amino acid region. Subsequent similarity searches using the *Volvox* AMC1-like protein yielded a *Gonium pectorale* hypothetical protein. The brown box depicts a 296 amino acid region with 46% similarity to *Gonium* hypothetical protein. The dark blue box depicts the region of low similarity (36%) between the *Chlamydomonas* AMC1 protein and the *Gonium* hypothetical protein. Comparison of these three proteins using the MEME program (BAILEY *et al.* 2009) identified the presence of two different motifs (motif 1 and motif 2, represented by a light blue box and a yellow box, respectively) that is common to the C-terminus of all three proteins. The sequence of these motifs in AMC1 are represented in Figure 7b.

(B) and **(C)** are the MEME-generated consensus sequences for motif 1 and motif 2, respectively (BAILEY *et al.* 2009). Motifs 1 and 2 are predicted to be 48 and 34 residues in length, respectively. The height of each letter represents the frequency of occurrence of the corresponding amino acid (y-axis) at a given position in the motif (x-axis). The *E-value* for the occurrence of motif 1 and motif 2 are $1.3e^{-006}$ and $1.8e^{-004}$, respectively.

SUPPORTING TABLES

Table S1. List of reference numbers for key *Chlamydomonas* strains used in this study.

Strain name	Genotype	Chlamydomonas collection center reference number
3A^{+}	*mt*$^{+}$; *arg7-8*	CC-5589
4C^{-}	*mt*$^{-}$; *arg7-8*	CC-5590
amc5 (87D3)	*mt*$^{+}$; *arg7-8*; *nuob10::APHVIII*	CC-5591
amc5; NUOB10	*mt*$^{+}$; *arg7-8*; *nuob10::APHVIII; NUOB10; ARG7*	CC-5592
amc1 (4C10)	*mt*$^{+}$; *arg7-8*; *amc1-1*; *APHVII*	CC-5597
amc1(2)	*mt*$^{+}$; *amc1-1*	CC-5598
amc1(27)	*mt*$^{-}$; *arg7-8; amc1-1*	CC-5599
amc11 (10G11)	*mt*$^{-}$; *arg7-8*; *amc1-2; APHVII*	CC-5608
amc1-2; AMC1-A	*mt*$^{-}$; *arg7-8*; *amc1-2; APHVII; AMC1; ARG7*	CC-5609
amc1-2; AMC1-B	*mt*$^{-}$; *arg7-8*; *amc1-2; APHVII; AMC1; ARG7*	CC-5610
amc11(65)	*mt*$^{-}$; *amc1-2; APHVII*	CC-5614

Table S1. List of reference numbers for key *Chlamydomonas* strains used in this study.

A list of the main strains used in this study is provided here with their genotypes and the reference number for retrieval from the *Chlamydomonas collection center* (chlamycollection.org).

Table S2. Sequence of primers.

Primer Name	Sequence (5' to 3')	Target
Cre16.g688900 5'UTR F1	GCACAGCCACACATGCATAAC	*AMC1* *(Cre16.g688900)*
10G11 exon2F (1)	GGTCAGGCTGGGATAAGTG	
10G11 AD1-F (2)	TACTGTTGCGGGTGCTG	
au5.g6830 exon2F2	TGCATAGCTCCTGCCTTCTC	
au5.g6830 exon2F3	AGGTGAGCTTCTGGATACG	
Cre16.g688900 exon2F4	CTGCGGAAAGGTGGCTTGAC	
Cre16.g688900 E2F32	TGGCGGGTGCCGCGCTGTTC	
Cre16.g688900 E2F33	CGCTGGTGCTGACGATAATG	
amc11 del1R	ATGACAGCGAGGGTCCCTTC	
Cre16.g688900 E2R9	GTGCTGCAAGACCGCTTTC	
Cre16.g688900 E2R10	TGTAGGACCGCCGGAGGCTTG	
au5.g6830 exon 3F	GCTACGGATGGCGAGGTG	
Cre16.g688900 exon3R2	GACGTGGCTTCCAGCTGTTG	
Cre16.g688900 exon4 F2	GGACACGTTGCAGAGGCAACAG	
au5.g6830 exon 4R	GTGCTGCGGAGTGTTATCG	
Cre16.g688900 exon4-3R	CAGGCACGCGGCCACCCCACGG	
Cre16.g688900 exon5 F3	CCAGCACCAGCATTCCATAG	
au5.g6830 exon6F2	CTGAGGGACACCAGTCTGTTC	
Cre16.g688900 exon8-1F	CTGTGAGCGGGATGAAAGG	
Cre16.g688900 exon8 R	AGGTGACTCGGCACCAAC	
Cre16.g688900 3'UTR-1R	AAGAGGTCACTGCCGTTCAC	
Cre16.g688900 3'UTR-2R	CCCGCATTGATGAACTTG	
PHA-F	CACACACGAATTCGATATCAAGCTGGAAGCTTAT GCGCGAGCATGCTGTGC.	

PHA-R	AAATACCGGCGATTTTTCGGCATTCATAAGCTTA ACCCCCTCGGGTGGCAGC.	
AD1	NTCASTWTSGWGTT	Partially degenerate primers used for TAIL-PCR
AD2	NGTCGASWGANAWGAA	
RMD228	WGNTCWGNCANGCG	
APH7R5	CGGTCGAGAAGTAACAGGG	iHyg3
APH7F3	CGACGTCTATGCGGGAGACT	
APH7F5	AACTGCATCTCAACGCCTTCC	
APH7F8	ACTGCTCGCCTTCACCTTC	
amc1-insert-R2	GGTACTCTTGACCCTAGTG	*amc1* insertion
amc1 insert F1	AGATAGCGTGAGATAGTACAG	
ND5-EG-F2	GATACCCTCCCAACCAACAAGCATAAC	*nd5*
cox1-R	CCGAATAGGGCTGGCATTAC	*cox1*
CCDA.13	GCCACATTCGCACTGGC	*CCDA*
CCDA.14	GGAGTCACCCAGGCCGAGTAC	

Table S2. Sequence of primers.

List of primers used for TAIL-PCRs, diagnostic PCRs, and RT-PCRs.

Table S3. Primers used for qRT-PCR.

Primer name	Sequence (5' to 3')	Target	Amplicon	Efficiency (%)	r^2 value
Chlamy cob fw	GCCTACCCAACTCCAATGAA	*cob*	192 bp	98.0	0.999
Chlamy cob rev	GTGAGCGTAACGCAAGATCA				
Chlamy nd4 fw	ACACTATGGCCGGTTCTTTG	*nd4*	213 bp	98.0	0.998
Chlamy nd4 rev	CACTACCAGCAGTTGGAGCA				
Chlamy nd5 fw	CCCCAATTGCTCGTTTTCTA	*nd5*	226 bp	97.0	0.997
Chlamy nd5 rev	CCGGTAACGGTGAATAGCAT				
Chlamy cox1 fw	TGGTAATGCCAGCCCTATTC	*cox1*	181 bp	98.0	0.999
Chlamy cox1 rev	TAAGCGGTCCAACCAGTACC				
Chlamy nd2 fw	CCACCATTTGCAGGTTTCTT	*nd2*	210 bp	100.0	0.999
Chlamy nd2 rev	GCAGGCAGAGGTTAGAGTGG				
Chlamy nd6 fw	TATTTTGTTGTGCGCTTTGC	*nd6*	220 bp	99.0	0.999
Chlamy nd6 rev	TAGCTCAGTGGCTGGGATCT				
ND1F2	GATCTACCAGAGGCTGAGTTG	*nd1*	150 bp	102.0	0.94
ND1R1	TTTAAGGGCGCTGAAGCCAC				
Chlamy rtl fw	CTGCCCTGCTTCTAATGGAG	*rtl*	231 bp	103.0	0.998
Chlamy rtl rev	TACCAAAACCAGGACGGAAG				

Cre16.g688900 exon5-1F	GGCCTGGCCGAGGAGATGGG	*AMC1*	82 bp	96.8	0.974
Cre16.g688900 exon5-2R	AGCAACTGGTGGGCTGTGAG				
EIFA FW	CATTGTGGAGCCGCCATTTC	*EIFA*	122 bp	99.8	0.997
EIFA REV	GGCTGCTTGCATTTGCTTCC				
CBLP-F	GCCACACCGAGTGGGTGTCGTGCG	*CBLP*	201 bp	108.0	0.9961
CBLP-R	CCTTGCCGCCCGAGGCGCACAGCG				
Tub-F	GTCCAAGCTGGGCTTCACCGTC	*TUA2*	152 bp	104.0	0.9908
Tub-R	GGCGGCAGATGTCGTAGATGGC				

Table S3. Primers used for qRT-PCR.

The sequences of primers used for qRT-PCR, with the resultant amplicon size, are shown here. The amplification efficiency for gene-specific primers, for the mitochondrial targets *cox1*, *nd4*, *nd5*, *cox1*, *nd2*, *nd6*, *nd1*, *rtl* (WOBBE AND NIXON 2013), and nuclear targets *CBLP* (ALLEN *et al.* 2007), *TUA2* (VALLEDOR *et al.* 2013), and *EIF1A* (SCHMOLLINGER *et al.* 2014) was identified by generating calibration curves with cDNA dilutions. The r^2 value for PCR efficiency, calculated from each calibration curve, are reported here. These PCR efficiencies were used in calculating the relative expression ratio by the Livak Method (LIVAK AND SCHMITTGEN 2001). The Phytozome IDs of nuclear genes are: *TUA2* - Cre04.g216850, *CBLP* - Cre13.g599400, *EIF1A* -Cre02.g103550.

SUPPORTING MATERIALS AND METHODS

Method S1. Genetic analysis

Genetic crosses were conducted according to (HARRIS 1989). Gametogenesis was induced by resuspending vegetative cells in TAP liquid medium lacking nitrogen (TAP-N), at 25° in low light (0.5-1.0 µmol. m^{-2}. s^{-1}), with shaking for 5 h. The gametes were mixed in equal proportions and incubated in light (50 µmol. m^{-2}. s^{-1}) at 25° overnight. To isolate meiotic zygotes, the mixture was plated on TAP-N solid medium (containing 3% (w/v) select agar) and incubated in light (50 µmol. m^{-2}. s^{-1}) at 25° for 5 days. The meiotic progeny was obtained through bulk germination on TARG solid medium, in continuous light at 25°. To isolate vegetative diploids, the mating mixture was directly plated on selective medium mentioned below. Individual diploids and haploid progeny were sub-cloned to a single colony on solid medium and their mating type was determined by diagnostic PCR (WERNER AND MERGENHAGEN 1998).

To determine the segregation of the SID phenotype in the *amc11* mutant, wild-type strain 1' (*mt*$^+$) [a *137C* derivative, kindly provided by Dr. Claire Remacle, University of Liège, Belgium] was mated with *amc11 (10G11)* (*mt*$^-$; *amc1-2*; *APHVII; arg7-8*). After bulk germination of the meiotic zygotes, haploid progeny were plated on TARG medium and the relevant phenotypes of the progeny were deduced by replica-plating. The haploid progeny derived from these crosses, *amc11 (65)* (*mt*$^-$; *amc1-2; APHVII),* was retained for further analyses.

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

Method S2. Genomic DNA extraction and diagnostic PCR

Genomic DNA was extracted from *Chlamydomonas* by phenol-chloroform method (SAMBROOK *et al.* 1989). The sequences of primers used in diagnostic PCRs and TAIL-

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

Method S3. TAIL-PCR (Thermal Asymmetric Interlaced PCR)

TAIL-PCR was used to identify the sequence flanking the insertional cassette in the *amc11 (10G11)* mutant as in (LIU *et al.* 1995). The following partially degenerate primers were used for TAIL-PCR: AD1, AD2, and RMD228 (LIU *et al.* 1995; DENT *et al.* 2005). The iHyg3-specific primers, APH7F3, APH7F5, and APH7F8, were used to amplify the genomic DNA flanking the cassette at its 3′-end. The site of insertion was identified as exon 2 of gene *Cre16.g688900*, as provided by the *Chlamydomonas* genome database version 5.5 in Phytozome version 12 (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Creinhardtii) (BLABY *et al.* 2014). The sequence flanking the 5′-end of the cassette was obtained by conventional PCR, using the *AMC1*-specific primer au5.g6830 exon2F3 and the iHyg3-specific primer APH7R5 (Figure S1).

To identify the *amc1* mutation, diagnostic PCR was conducted and an insertional mutation was identified in exon 2. TAIL-PCR was used to identify the insertional sequence disrupting the *AMC1* gene in the *amc1-1* mutant by analyzing two independent *amc1-1* strains: *amc1 (4C10)* and *amc1 (27).* Two sets of TAIL-PCR reactions were successful. In both, the partially degenerate primer RMD228 was used, in combination with the following *AMC1*-specific primers (for primary, secondary, and tertiary reactions, respectively): A) Cre16.g688900 E2F32, Cre16.g688900 E2F33, and au5.g6830 exon2F2, and B) Cre16.g688900 E2F33, au5.g6830 exon2F2, and 10G11 exon2F (1). The complete sequence was retrieved by amplification with the primer pair amc1insertF1/amc11del-R. The inserted DNA, disrupting the *AMC1* gene, was confirmed

by sequencing two independent *amc1-1* strains: the original *amc1 (4C10)* and its haploid progeny *amc1(27)* (Figure S2).

Method S4. PCR-based screening of *Chlamydomonas* genomic library

An *ARG7*-based indexed cosmid library of *Chlamydomonas* genomic DNA, created by Dr. Jean David Rochaix's laboratory (University of Geneva, Switzerland) (PURTON AND ROCHAIX 1994), was screened for the presence of cosmids carrying the *AMC1* gene by diagnostic PCR. The *AMC1*-containing cosmid (referred to as 1H10) was identified using the primer pair au5.g6830 exon2F3/10G11 AD1-F (2). In order to generate a construct containing only the full-length *AMC1* gene, the cosmid 1H10 was digested with *Bam*H1 restriction enzyme and recircularized by T4 DNA ligase (Invitrogen, 15224041). The borders of *Chlamydomonas* genomic DNA insert, present in the original cosmid 1H10 (referred to as cosmid A) and the *Bam*HI digested and re-ligated cosmid (referred to as cosmid B) were sequenced to confirm the presence of the region carrying the gene of interest (Figure S3A).

Method S5. Biolistic transformation

Cosmids A and B were used for transformation by biolistics. The *ARG7* gene in the cosmids was used as a selection marker. The recipient arginine auxotrophic strain *amc11 (10G11) (mt*$^-$; *amc1-2*; *APHVII*; *arg7-8)*, also referred to as the *amc1-2* mutant in this manuscript, was subjected to biolistic transformation using a homemade particle delivery device. The strains were grown in liquid TARG medium for 2-3 days until they reached exponential phase (3 – 6 x 10^6 cells. mL^{-1}). The cells were plated on selective TAP medium at 10^8 cells/plate. For each bombardment, DNA was coated on sterile 0.6-0.9 µm tungsten particles (STREM Chemicals, 93-7437) by using 2 µg of cosmid, 16.7 mM Spermidine (Sigma, S2626-5G), and 1 M $CaCl_2$. The bombardment was conducted at a helium pressure of 1.725 MPa and vacuum of -92 kPa. The plate was positioned 10.5 cm away from the nozzle containing the coated particles. The bombarded plates were first incubated at low light overnight for recovery and then transferred to continuous light (50 µmol. m^2. s^{-1}).

Transformants from cosmid A were screened by replica-plating. Out of 311 transformants tested, 115 transformants displayed rescued growth in the dark. Transformants from

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

Method S6. RNA extraction, RT-PCR, and qPCR

To determine the *AMC1* gene model (Figure S4), RNA was prepared using the Plant RNeasy Kit (Qiagen, 74904) with the following modifications. Cells were grown for 2-3 days on TARG solid medium. Cells (100 mg) were harvested and lysed by vortexing, with 9/10th volume of glass beads, for 5 min at 22º. RNA extraction was completed according to manufacturer's protocols. WT (4C^{-}) RNA (2-5 µg) was treated with RQ1 RNase-free DNase I (Promega, M6101). Reverse-transcription was achieved with 400 units of M-MLV reverse transcriptase (Life Technologies, 28025-013). Some regions of the cDNA could be amplified only if the first strand synthesis was conducted with Roche Transcriptor High Fidelity cDNA synthesis kit (05081955001), using OligodT primers or random hexamers as per manufacturer's protocol. Different regions of the *AMC1* cDNA were successfully amplified using either GoTaq Polymerase (Promega, M3008) or Phusion High-Fidelity DNA Polymerase (NEB, B0519S). The primer pairs used for successfully sequenced cDNA amplicons are provided in the legend of Figure S4. The amplicon generated using Cre16.g688900 5′UTR-F1/E2R10 was cloned into Promega pGEM-T easy Vector System. The resulting construct contains the sequence corresponding to the AMC1 N-terminal that was cloned in-frame to the *ubiG* reporter.

For real-time quantitative PCR (qPCR) of the mitochondrial transcripts, RNA was isolated by phenol-chloroform method modified from (NEWMAN *et al.* 1990). RNA was extracted from 2 x 10^{8} cells grown in liquid culture. The cells were first resuspended in TEN buffer (10 mM Tris-HCl pH 8.0, 10 mM EDTA, 150 mM NaCl) and pelleted. The cells were resuspended in 150 µl of water and 300 µl of SDS-EB buffer (2% SDS, 400 mM NaCl, 40 mM EDTA, 100 mM Tris-HCl pH 8.0). Nucleic acids were extracted twice with equal volume of phenol-chloroform (pH 5.0) and RNA was precipitated with 1/3rd volume of 8 M LiCl. RNA (8 µg) was treated with RQ1 RNase-free DNase I (Promega, M6101). Reverse

transcription was achieved with 800 units of M-MLV Reverse transcriptase (Life Technologies, 28025-013) using 1 µg of Random Hexamers (Promega, C1181), according to the manufacturer's protocol. The absence of contaminating mtDNA was confirmed by diagnostic PCR of cDNA sample, across the two transcription start sites, with the primer pair ND5-EG-F2 and cox1-R (Table S2). The qPCR was conducted simultaneously on two dilutions of the cDNA and the results were compared for correlation. Amount of cDNA equivalent to 50 ng or 100 ng of total input RNA was used as template for qPCR using SensiMix (Bioline, QT-650-05) on a Mastercycler ep gradientS realplex thermocycler (Eppendorf). The qPCR reactions were denatured at 98º for 15 s, annealed at 60º for 20 s and extended at 72º for 20 s. Relative transcript levels were determined by normalizing the levels of target transcripts to the geometric mean of three reference transcripts, *TUA2, CBLP*, and *EIF1A.* For determining the relative mitochondrial DNA (mtDNA) content, 8 ng of total genomic DNA was used as template for qPCR. The mitochondrial *nd4* gene was used as the target gene and the nuclear *TUA2* gene was used as the reference. In all cases, normalization of target transcript or gene levels to the reference was performed by the Livak $2^{-\Delta\Delta Ct}$ method (LIVAK AND SCHMITTGEN 2001). Relative fold change for mtDNA content was calculated by normalizing to the average of the isogenic wild-type strain (set to 1.0). Relative transcript abundance is represented as percentage of wild-type strain. The amplification efficiency for each primer pair was determined by generating calibration curves for each pair of primers using two-fold dilutions of the template (wild-type cDNA). The efficiency was calculated from the slope of the calibration line as $[10^{(-1/slope)}-1] \times 100$. The efficiency and r^2 value from the calibration line are reported in Table S3. In Figure 6C, the *amc1-1* and *amc1-2* mtDNA content were normalized to their respective WT, $3A^+$ and $4C^-$, respectively. The strains tested in Figure 6D are WT ($4C^-$), *amc1-2* [*amc11 (10G11*)], and [*amc1-2; AMC1-B*]. The strains tested in Figure 6E are WT ($3A^+$) and *amc1-1* [*amc1(4C10*)].

Method S7. Plasmid construction to test for mitochondrial localization

The AMC1 N-terminal mitochondrial targeting sequence was heterologously expressed as a translational fusion with the bacterial UbiG protein in *S. cerevisiae*. For this purpose, plasmids pAHG (*CYC1* promoter + *ubiG*), pQMG (*CYC1* promoter + *ubiG* with 5'-end sequence encoding the COQ3 N-terminal mitochondrial targeting signal), and pAH3

(*CYC1* promoter + *COQ3*), provided by Dr. Catherine Clarke (University of California, Los Angeles) (HSU *et al.* 1996), were used. The sequence encoding the AMC1 N-terminus (1-59 amino acids) was amplified, using the primers PHA-F and PHA-R, from a plasmid containing a cloned fragment of the *AMC1* cDNA (see Method S6). The PCR product was cloned into the *Hin*dIII site of the plasmid pAHG (HSU *et al.* 1996), carrying the *ubiG* gene under the control of the constitutive *CYC1* promoter, by In-Fusion Cloning (In-Fusion HD cloning kit, Clontech, 639648). The resulting construct, containing the sequence corresponding to 1-59 amino acids of AMC1 at the 5'-end of the *ubiG* gene, is referred to as pAHG59. The pAHG, pQMG, pAH3, and pAHG59 plasmids, each containing the *URA3* selectable marker, were introduced into the *Δcoq3* strain by one-step transformation (CHEN *et al.* 1992).

SUPPLEMENTAL INFORMATION REFERENCES

Allen, M. D., J. Kropat, S. Tottey, J. A. Del Campo and S. S. Merchant, 2007 Manganese deficiency in Chlamydomonas results in loss of photosystem II and MnSOD function, sensitivity to peroxides, and secondary phosphorus and iron deficiency. Plant Physiol 143**:** 263-277.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman, 1990 Basic local alignment search tool. J Mol Biol 215**:** 403-410.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant *et al.*, 2009 MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37**:** W202-208.

Blaby, I. K., C. E. Blaby-Haas, N. Tourasse, E. F. Hom, D. Lopez *et al.*, 2014 The Chlamydomonas genome project: a decade on. Trends Plant Sci 19**:** 672-680.

Boratyn, G. M., A. A. Schaffer, R. Agarwala, S. F. Altschul, D. J. Lipman *et al.*, 2012 Domain enhanced lookup time accelerated BLAST. Biol Direct 7**:** 12.

Chen, D. C., B. C. Yang and T. T. Kuo, 1992 One-step transformation of yeast in stationary phase. Curr Genet 21**:** 83-84.

Dent, R. M., C. M. Haglund, B. L. Chin, M. C. Kobayashi and K. K. Niyogi, 2005 Functional genomics of eukaryotic photosynthesis using insertional mutagenesis of *Chlamydomonas reinhardtii*. Plant Physiol 137**:** 545-556.

Gabilly, S. T., J. Kropat, M. Karamoko, M. D. Page, S. S. Nakamoto *et al.*, 2011 A novel component of the disulfide-reducing pathway required for cytochrome c assembly in plastids. Genetics 187**:** 793-802.

Harris, E. H., 1989 *The Chlamydomonas Sourcebook: A comprehensive guide to biology and laboratory use*. Academic Press, San Diego, CA.

Hsu, A. Y., W. W. Poon, J. A. Shepherd, D. C. Myles and C. F. Clarke, 1996 Complementation of coq3 mutant yeast by mitochondrial targeting of the Escherichia coli UbiG polypeptide: evidence that UbiG catalyzes both O-methylation steps in ubiquinone biosynthesis. Biochemistry 35**:** 9797-9806.

Liu, Y. G., N. Mitsukawa, T. Oosumi and R. F. Whittier, 1995 Efficient isolation and mapping of *Arabidopsis thaliana* T-DNA insert junctions by thermal asymmetric interlaced PCR. Plant J. 8**:** 457-463.

Livak, K. J., and T. D. Schmittgen, 2001 Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25**:** 402-408.

Newman, S. M., J. E. Boynton, N. W. Gillham, B. L. Randolph-Anderson, A. M. Johnson *et al.*, 1990 Transformation of chloroplast ribosomal RNA genes in Chlamydomonas: molecular and genetic characterization of integration events. Genetics 126**:** 875-888.

Purton, S., and J. D. Rochaix, 1994 Complementation of a *Chlamydomonas reinhardtii* mutant using a genomic cosmid library. Plant Mol. Biol. 24**:** 533-537.

Sambrook, J., E. F. Fritsch and T. Maniatis, 1989 *Molecular Cloning A Laboratory Manual*. Cold Spring Harbor Laboratory Press, New York.

Schmollinger, S., T. Muhlhaus, N. R. Boyle, I. K. Blaby, D. Casero *et al.*, 2014 Nitrogen-Sparing Mechanisms in Chlamydomonas Affect the Transcriptome, the Proteome, and Photosynthetic Metabolism. Plant Cell 26**:** 1410-1435.

Valledor, L., T. Furuhashi, A. M. Hanak and W. Weckwerth, 2013 Systemic cold stress adaptation of Chlamydomonas reinhardtii. Mol Cell Proteomics 12**:** 2032-2047.

Werner, R., and D. Mergenhagen, 1998 Mating Type Determination of Chlamydomonas reinhardtii by PCR. Plant Molecular Biology Reporter 16**:** 295-299.

Wobbe, L., and P. J. Nixon, 2013 The mTERF protein MOC1 terminates mitochondrial DNA transcription in the unicellular green alga Chlamydomonas reinhardtii. Nucleic Acids Res 41**:** 6553-6567.