

Small RNA profiling in *Chlamydomonas*: insights into chloroplast RNA metabolism.

Marina Cavaiuolo, Richard Kuras, Francis-André Wollman, Yves Choquet, Olivier Vallon

Supplementary Figure S1. A stable stem-loop structure at *psaJ* 3'UTR.

Supplementary Figure S2. Characteristics of Cp small RNAs

Supplementary Figure S3. Determination of the 5' and 3' boundaries of tRNA precursors

Supplementary Figure S4. WTSS reveals co-transcription of consecutive genes on the same strand.

Supplementary Figure S5. Size distribution of sRNAs mapping to *psbA* upon transcriptional and translation inhibition.

Supplementary Figure S6. Distribution and relative abundance of RPFs after CAP treatment.

Supplementary Figure S7. Size distribution of 11 to 44-nt antisense sRNAs in Linc, CAP and Rif treated samples.

Supplementary Table S1. Directional WTSS and sRNA-Seq datasets produced in this work and their mapping statistics.

Supplementary Table S2. Bi-directional WTSS libraries available from NBCI and used in this work.

Supplementary Table S3. Primers used for cRT-PCR, 5'RACE, strand-specific RT-PCR, qPCR and verification of insertion mutants.

Supplementary Table S4. 5' cosRNA in relationship to chloroplast genes and gene clusters.

Supplementary Table S5. Predicted secondary structures in the 3'UTRs of Cp genes.

Supplementary Table S6. Tentative assignment of 5' cosRNA to genetically identified M factors.

Supplementary Table S7. Effect of translational inhibition on sRNAs and RNA levels of *psbA*.

Supplementary Table S8. Effect of transcription inhibition on sRNAs and RNA levels of *psbA*.

Supplementary Table S9. Differential expression analysis: WTSS and sRNA-Seq coverage of protein-coding genes (sense strand).

Supplementary Table S10. RPM of all sRNA-Seq reads, RPFs and non-RPFs mapped to CDS of Cp protein coding genes after CAP and Rif treatment in WT strains.

Supplementary Table S11. Differential expression analysis for protein-coding genes: WTSS coverage between mixo- and photo-trophic conditions (sense strand).

Supplementary Table S12. Mean expression levels of Cp genes in mixo- or photo-trophic conditions.

Supplementary Table S13. Differential expression analysis: WTSS coverage for nuclear genes.

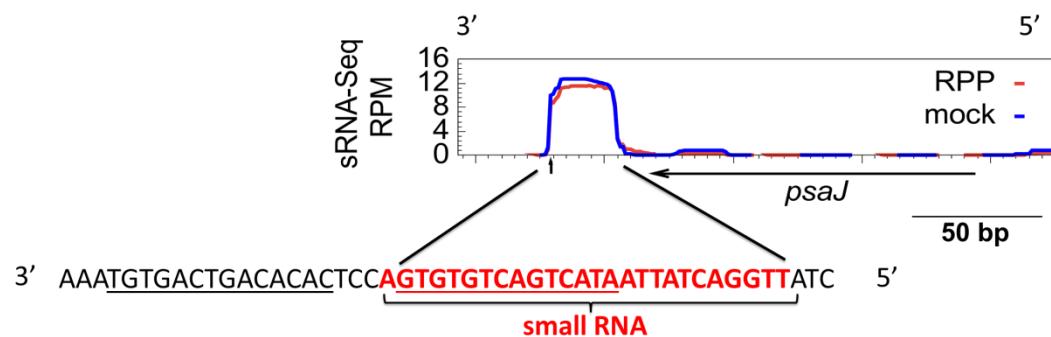
Supplementary Table S14. Comparison of WTSS and sRNA-Seq coverage on sense and antisense strands for Cp regions.

Supplementary Table S15. Differential expression analysis: sRNA-Seq coverage of protein-coding genes (antisense strand).

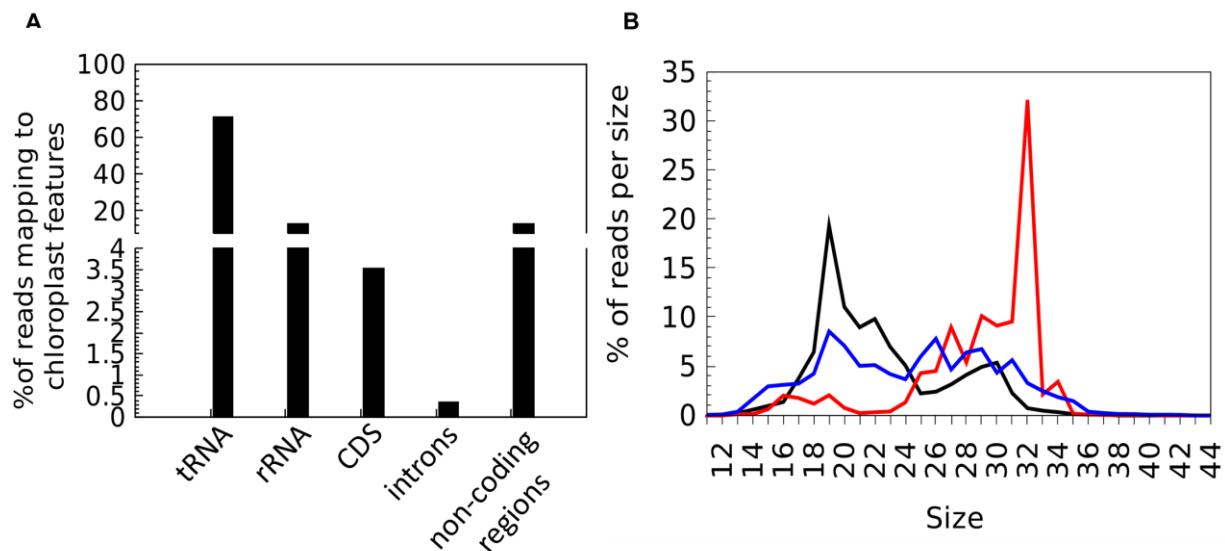
Supplementary Table S16. sRNA-Seq results of CAP-treated WT cc-4533 and mutants in TCA1, PPR and OPR genes (sense strand).

Supplementary references. References cited in the supplementary Figures and Tables.

Supplementary Figure S1. A stable stem-loop structure at *psaJ* 3'UTR. sRNA-Seq profile downstream of *psaJ* reveals a RPP-independent cosRNA, written in red, overlapping with a secondary structure. The cosRNA lies within the first arm of an inverted repeat (underlined) ($\Delta G = -25.2$, Mfold).



Supplementary Figure S2. Characteristics of Cp small RNAs. A) Proportion of sRNA-Seq reads that map to either strand of tRNAs, rRNAs, CDS, introns or non-coding regions (UTRs and intercistronic/intergenic regions). B) Size distribution of sRNA-Seq reads mapping to tRNA (red), rRNA (blue) and other regions (CDS + non-coding regions; black).

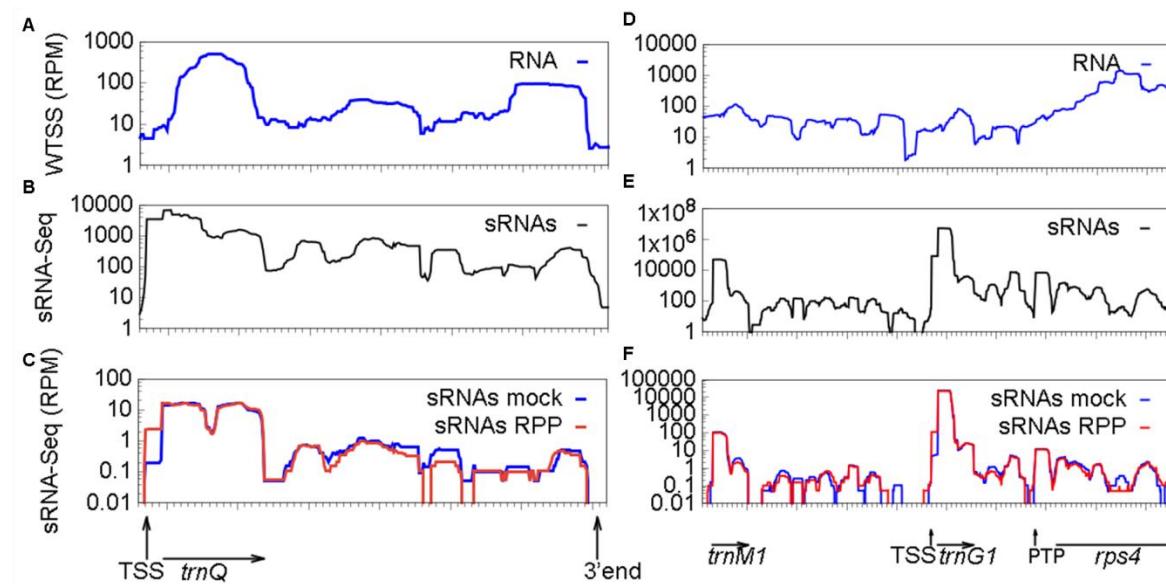


Supplementary Figure S3. Determination of the 5' and 3' boundaries of tRNAs precursors. The figure displays the genomic region encompassing the *trnQ* (A-B-C) or *trnM1* and the co-transcribed *trnG1* and *rps4* genes (D-E-F).

A, D) WTSS coverage in pooled directional and bi-directional datasets. In D, a low but continuous coverage extends in the *trnM1-trnG1* intergenic regions.

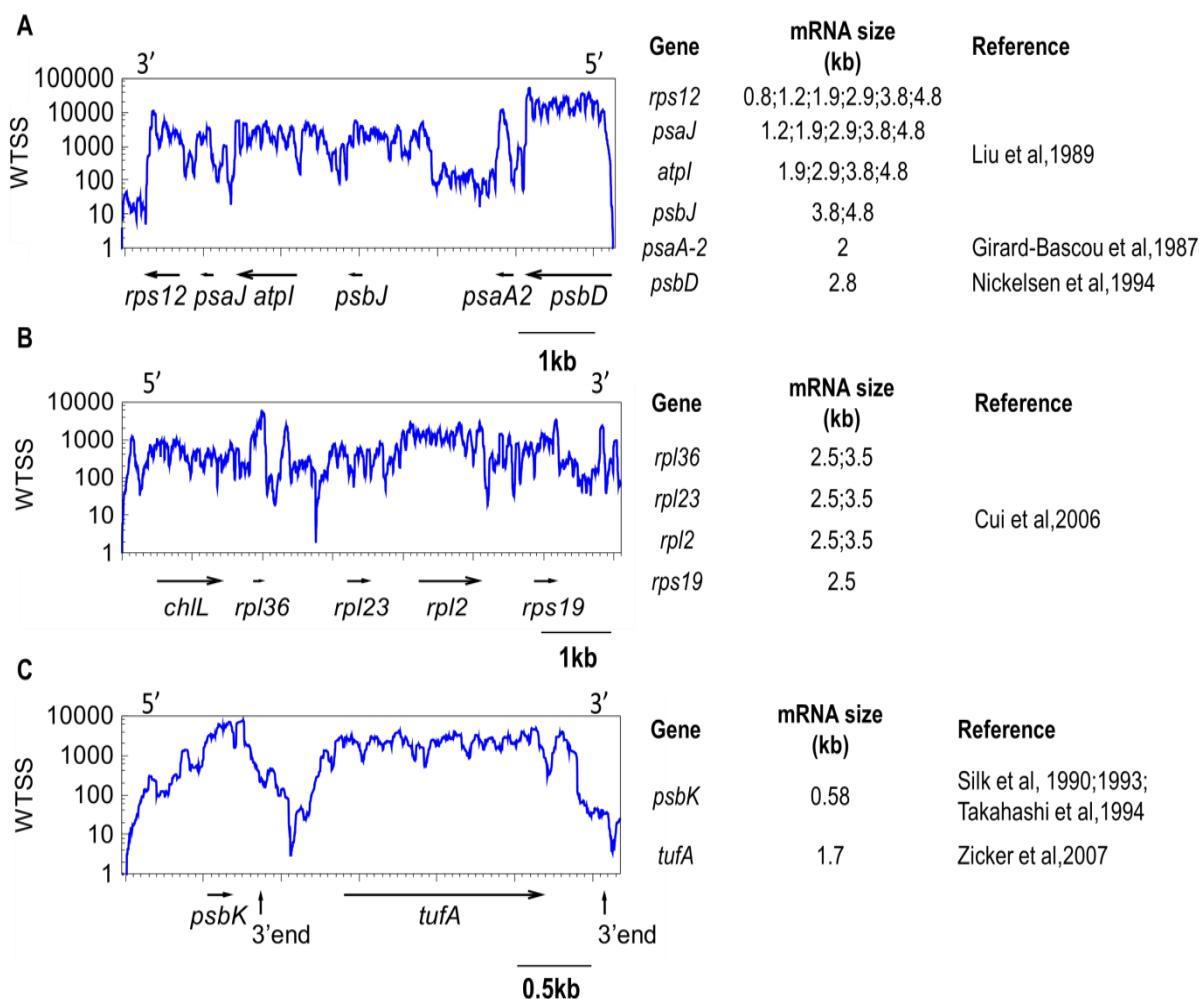
B, E) sRNA-Seq coverage (perfect match) from the pooled datasets shows the sRNAs produced from processing and degradation of mature and precursors tRNAs. Two major fragments are produced, one starting at the 5' end (a possible RNase P processing site) of the mature tRNA and one ending at the 3' end of the tRNA (a possible RNase Z processing site). Other sRNA mapping upstream and downstream were used to define tRNA precursor boundaries. A coverage cutoff of 15 reads was set.

C, F) cosRNAs upstream of tRNA correspond to a TSS, as judged from higher coverage in RPP- (in red) compared to mock-treated samples (in blue). Coverage was averaged between 2 biological replicates and normalized as reads per million (RPM).

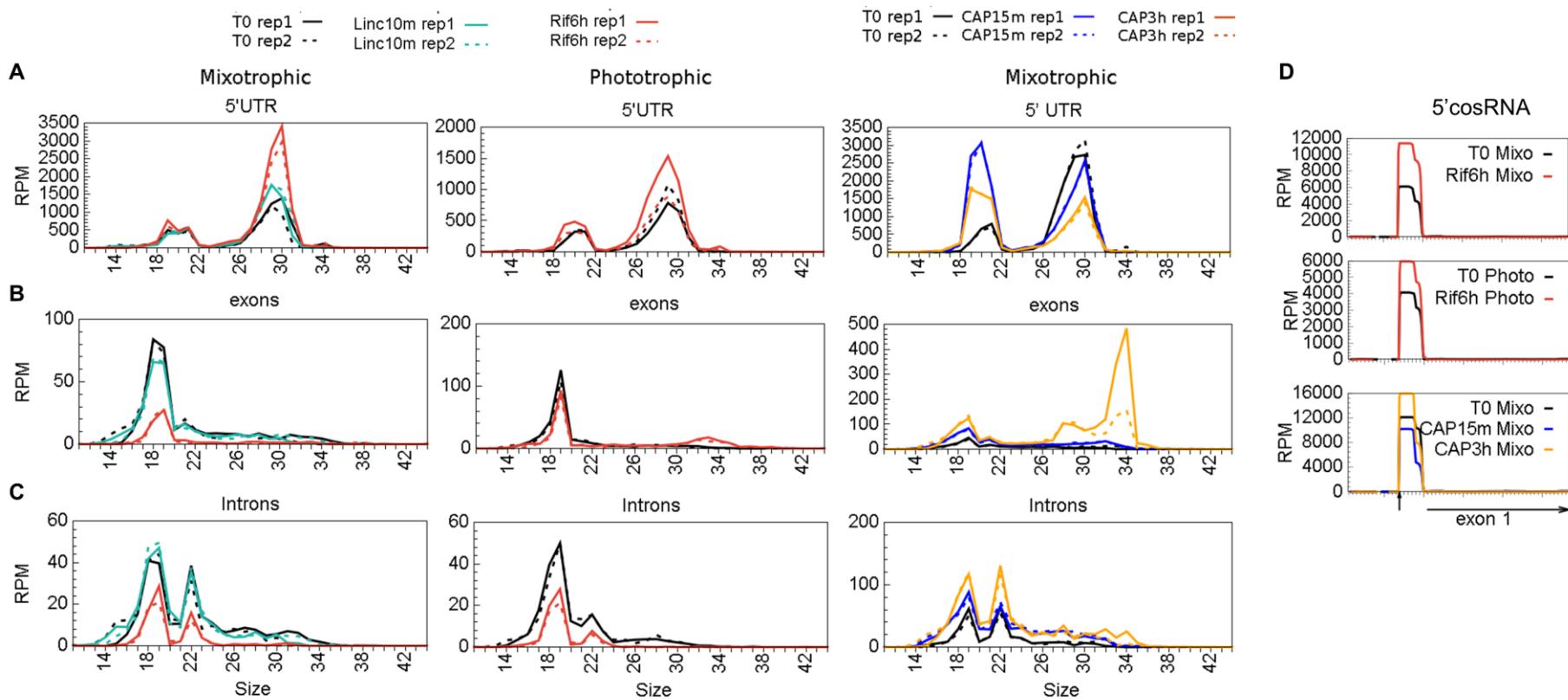


Supplementary Figure S4. WTSS coverage reveals co-transcription of consecutive genes on the same strand. Total read counts from bi-directional and directional datasets. Horizontal arrows indicate CDS. The sizes of the major transcripts previously detected by RNA blots are shown along.

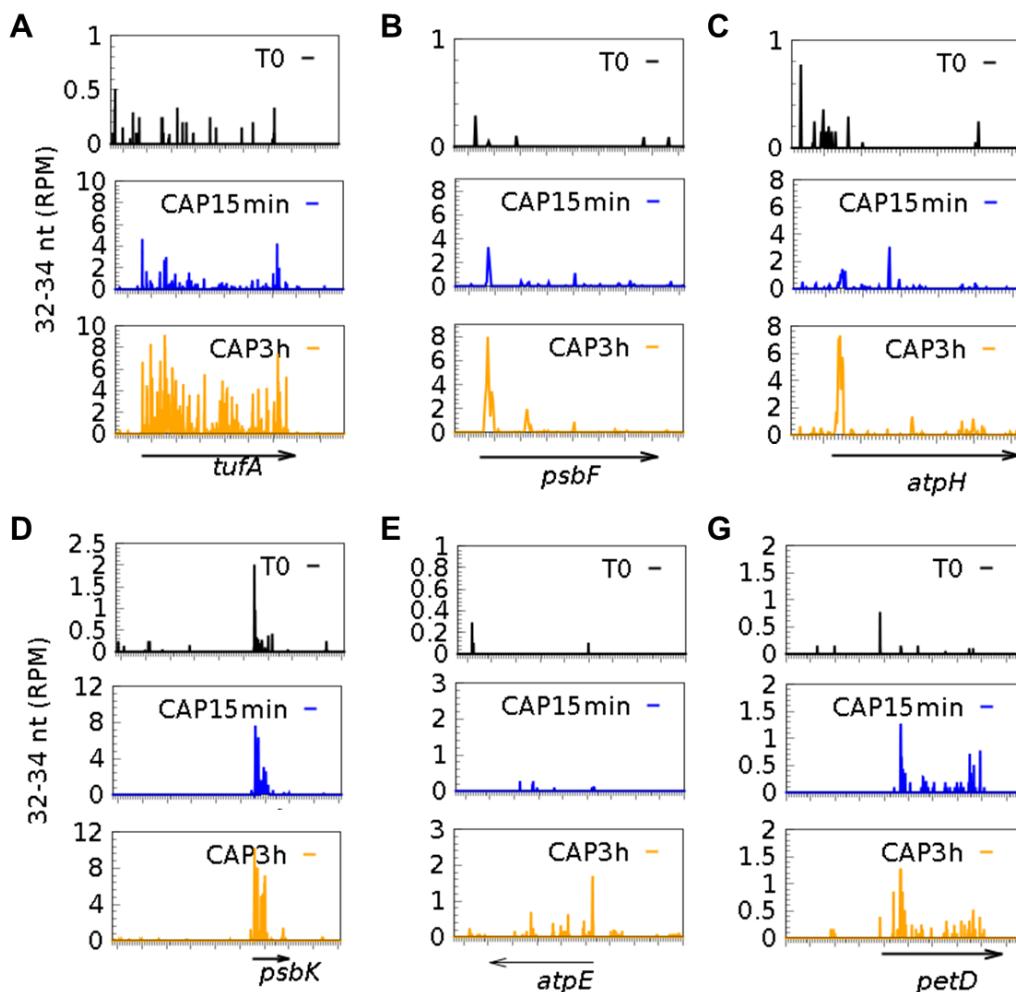
- A) Continuous coverage indicates co-transcription of the *psbD-psaA.2-psbJ-atpl-psaJ-rps12* gene cluster.
- B) Near-continuous coverage indicates co-transcription of the *chIL-rpl36-rpl23-rpl2-rps19* genes. Bands corresponding to bi-cistronic *rpl36-rpl23* and *rpl2-rps19* and to tricistronic *rpl36-rpl23-rpl2* transcripts were reported earlier, but none of a size consistent with a long polycistronic transcript encompassing the five genes. The coverage dip between *rpl36* and *rpl23* does not match a cosRNA nor a promoter sequence. Note that transcription extends further downstream to cover the *rpl16-rpl14-rpl5-rps8* genes.
- C) Extension of coverage downstream of the mature 3'ends for two monocistronic mRNA *psbK* and *tufA* transcripts (vertical arrows) suggests a possible co-transcription of the two genes, even if *tufA* can be transcribed independently from its own promoter.



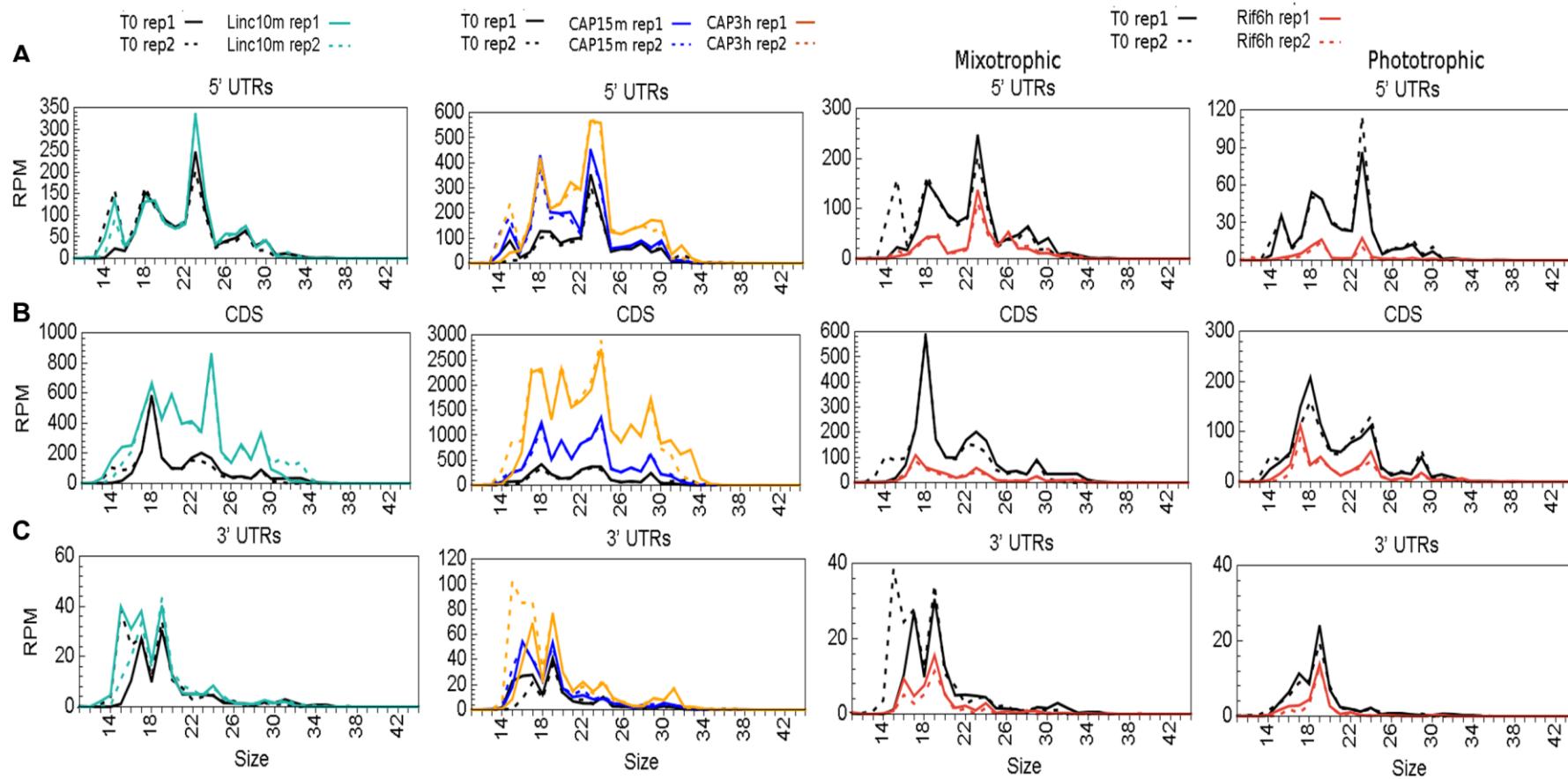
Supplementary Figure S5. Size distribution of sRNAs mapping to *psbA* upon transcriptional and translation inhibition. Abundance of sRNAs is plotted as a function of their length for sRNAs mapping to 5'UTR (A), exons (B) and introns (C). Two major populations of sRNAs, 19-21 nt and 28-30 nt, originate from the 5'UTR of *psbA* (A) and co-localize with the mature 5'ends of *psbA* (D). RPM were normalized on the total mapped reads of nuclear and mitochondrial genomes. For each experiment, two replicates are shown to illustrate the reproducibility of the results. Read mapping to CDS encoding maturase-like proteins were subtracted from those of introns.



Supplementary Figure S6. Distribution and relative abundance of RPFs after CAP treatment.
The profiles of 5' end positions of 32-34 nt sRNA sequences in *tufA* (A), *psbF* (B), *atpH* (C), *psbK* (D), *atpE* (E) and *petD* (G) illustrate the prevalence of RPFs in CDS (indicated by the horizontal arrow). Values are expressed as RPM, normalized to the total mapped reads of nuclear and mitochondrial genomes and averaged between two replicates.



Supplementary Figure S7. Size distribution of 11 to 44-nt antisense sRNAs in Linc, CAP and Rif treated samples. Abundance of sRNAs, expressed as RPM normalized to the number of reads mapping to the nuclear and mitochondrial genomes, is plotted as a function of sRNA length for 5'UTRs (A), CDS (including *psaA* and *psbA*) (B) and 3'UTRs (C).



Supplementary Table S1. Directional WTSS and sRNA-Seq datasets produced in this work and their mapping statistics. (see “supplementary_data_1.xlsx”, attached)

Supplementary Table S2. Bi-directional WTSS libraries available from NBCI and used in this work. (see “supplementary_data_1.xlsx”, attached)

Supplementary Table S3. Primers used for PCR, cRT-PCR, 5'RACE, strand-specific RT-PCR and qPCR.

Name	Sequence	Application	Target
psbK_cRT	TTGGTTTCTTCCCACAAACC	cRT-PCR	<i>psbK</i>
psbK_F1	GCAGTTGACAGGCTTAGACC	cRT-PCR	
psbK_R1	TTCGCGTACCCGTAAAGAGT	cRT-PCR	
psbF_cRT	AGGGACGAATACTTAACGTTGA	cRT-PCR	<i>psbF</i>
psbF_cr2	CCAACGAACTGTGAAAATAGGA	cRT-PCR	
psbF_cf2	ATATTCAAAGGGCGGCTGT	cRT-PCR	
petB_cRT	TTCTGCTACTGTTGGACGGT	cRT-PCR	<i>petB</i>
petB_cf2	AGCTACTTGCACAAGGAAACA	cRT-PCR	
petB_cr1	CAGGTGGTTCAAAACGTCCA	cRT-PCR	
atpE_cRT	TCAACACCTGCTCTACTCTGT	cRT-PCR	<i>atpE</i>
atpE_cf1	TCACACTAGAACCTTTGTCGG	cRT-PCR	
atpE_cr1	GGTGGTCAAGCATCAGGTT	cRT-PCR	
petA_cRT	GCGATCTGGTGGTGCTAATT	cRT-PCR	<i>petA</i>
petA_cf1	AGAAGGTCAAACACTGTACAAGCAGATCA	cRT-PCR	
petA_cr1	AAAGTAGTAAATACTGGTTAGAC	cRT-PCR	
atpI_R1	TTTGGAAACCATCAGGAGTT	strand-specific RT	<i>antisense atpI</i>
atpI_F1	TAACCTTACTTACGTGTCCCTC	strand-specific RT	
petA_F1	GTATGGTAGGGTTATCC	strand-specific RT	
petA_F2	TGTATGTCAAACGTCACT	strand-specific RT	<i>antisense petA</i>
petA_F3	AGCAATCACAGCTTTCTG	strand-specific RT	
petA_R1	CTACAACAACTTCACCGTT	strand-specific RT	
petA_5'RACE	CACCAGATCGCGTCCGGCAGAAATTAA	5'RACE	
atpA_F1	GACAAGGGTGAACCATTACT	strand-specific RT	<i>antisense atpA</i>
atpA_F2	GTCCAGTTGATGGTAAAGGT	strand-specific RT	
atpA_F3	CAGCTCAACCAAAAGCAATG	strand-specific RT	
atpA_R1	TAAGCAGCTTAGCTTGAGA	strand-specific RT	
atpB_F1	TGGCTTTAAGAAGAAAACA	strand-specific RT	<i>antisense atpB</i>
atpB_R1	CAGTTAAAGGTGCCCTTTGA	strand-specific RT	
atpB_F	ACGAGCACGAGCTACAATAAGT	strand-specific qPCR	
atpB_R	CCTGCGTTGACCCATTAGA	strand-specific qPCR	
16SrRNA_F	ATGGAGACTAAGTGTGCCG	strand-specific qPCR	<i>16S</i>
16SrRNA_R	CCTGGTAAGGTTCTCGCGT	strand-specific qPCR	
PPR1-F	TCCATGAGCATAACGGTACGA	PCR	PPR1
PPR1-R	CACAATACATTGCCAGTG	PCR	
PPR3-F	TCACCTACAGTGCCTCATC	PCR	PPR3

PPR3-R	CGGTACCACCAAGCAACAATA	PCR	
PPR6-F	ACGAGTAGCTGCCAAAGGAA	PCR	PPR6
PPR6-R	GGCTGGGATAGTTAGGAGGC	PCR	
OPR56-F	GCTGTCCATGCTGATGTACG	PCR	OPR56
OPR56-R	CAAGCATAGCCGGACTGAAC	PCR	
OPR105-F	CTTCTGAAACTGGAAGCGGT	PCR	OPR105
OPR105-R	TCGATTAGTTGGGCTGGT	PCR	
OPR49-F	CTACCACCCCCAGCTGTG	PCR	OPR49
OPR49-R	GCTCTGCACCAAATGAGACA	PCR	
OPR41-F	AATCGATATGTCCCCTTCC	PCR	OPR41
OPR41-R	TGTGCATAGTGTAGGCGCTC	PCR	
OPR24-F	AGAATATATCAGGCGCCGTC	PCR	OPR24
OPR24-R	GGTGTGAAGGAGGCCATGT	PCR	
oMJ282-F	ATGCTTCTCTGCATCCGTCT	PCR	oMJ282
oMJ282-R	ATGTTTACGTCCAGTCCGC	PCR	

Supplementary Table S4. 5' cosRNA in relationship to chloroplast genes and gene clusters.

(see “supplementary_data_1.xlsx”, attached)

Supplementary Table S5: Predicted secondary structures in the 3'UTRs of Cp genes. The genomic position of the start and end of the secondary structure is shown.

The sequence of mapped sRNA read is written in bold. The 3'-end nucleotides are shown in red, in bold if identified by cRT-PCR, italicized if collected from the literature. The minimum free energy (MFE) (kcal/mol) was predicted using Mfold. The asterisk indicates that the sequence is found in other parts of the genome and the sRNA is not reported because the true origin cannot be ascertained.. Only a unique sRNA sequence found in conjunction with a predicted stem-loop was used to annotate the 3'-ends.

Gene (strand)	Start	End	Potential secondary structure in 3'UTR	MFE
<i>wendyA*</i> (-)	203994	203918	UUAAUACUGCGGAGC AGGCAGUGGCGGUACCACA AUAAAUCAA <u>UUUGUCCUGCCAACUGCCU</u> GCUCGGCAGUAUUA (((((((((((((.....((((.....))))..)))))))))))))))))))	-48.7
<i>petA</i> (+)	3840	3907	AAAUGAACUUCAAUAUUUAUUUUUUGUAGGGCU GCUGUGCAGCU CUACAAAUUUAGUAUGUUA(((((.....((((((.....))))..))))))))..))))	-20.9
<i>petD</i> (+)	6896	6939	UUUCCUCUAGGGUUGCAAUACGAUUUGCAACCCUGAAGGGGGAAAACUGAG U (((((((((((((.....))))))))..)))))))).....	-29.0
<i>rpl20</i> (-)	16053	16025	AACAGUUGGUGGUACUACCAACUGCU ..(((((((((.....))))))))..))	-18.6
<i>cIPP</i> (-)	17773	17729	UUUUUAUAAAAGGGUAUCACUACAAGGUAUCCCUUAAAUA (((((((((((((.....))))))))))))))))	-22.2
<i>petB</i> (-)	20150	20096	AUUGAAGUUUAUUUACCAAAGGGAUACACCCUUUGGGUA AUAAACU U CAAU (((((((((((((.....))))))))))))))))))))	-39.9
<i>rps19</i> (+)	27218	27268	AAGAACUUGUUGUCUUGCAGCUC UUUGCGCGCUGCCAGACGGCAAGUUCU ((((((((((.(((.....))))..))))))))))))	-29.6
<i>rpl16</i> (+)	28972	29008	GAAAAUUUAUUUUUUCACUUAUGAAAAAAUAAA ((((((((((.....))))))))))))	-11.9
<i>rpl14*</i> (+)	29674	29706	GAA GGAGGCAGUUGGCAGGCAACUG CCUCCUUC ((((((((((.....)))))))))))	-25.1

Supplementary Table S6. Tentative assignment of 5' cosRNA to genetically identified M factors.

Gene	Strand	5' of cosRNA	proposed M-factor :	Repeat type	subjected to mutagenesis	Reference
<i>atpE</i>	-	61465 (PTP)	MDE1	-	-	Drapier et al., unpublished
<i>petA</i>	+	2640 (TSS)	MCA1	PPR	2640-2661	Loiselay et al , 2008
<i>petD</i>	+	6038 (PTP)	MCD1	OPR	6038-6062	Higgs et al, 1999; Drager et al, 1998,1999; Murakami et al 2005
<i>petG</i>	+	104047 (PTP)	MCG1	OPR	-	Wang et al, 2015
<i>psbA.2</i>	+	139619 (PTP)	RBP63	-	139619-139654	Ossenbühl et al, 2002
<i>psbB</i>	-	82553 (PTP)	MBB1	TPR	82533-82553	Loizeau et al, 2012
<i>psbH</i>	-	77778 (PTP)	MBB1	TPR	77758-77778	Loizeau et al, 2012
<i>psbC</i>	+	188039 (PTP)	MBC1(OPR56)	OPR	-	This study
<i>psbD</i>	-	177235 (PTP)	MBD1 (NAC2)	TPR	177209-177235	Nickelsen et al, 1994, 1999; Bruick and Mayfield, 1998
<i>psbI</i>	+	127439 (PTP)	MBI1	OPR	-	Drapier et al, 1998; Wang et al, 2015

Supplementary Table S7. Effect of translational inhibition on sRNAs and RNA levels of *psbA*. sRNA and RNA levels are reported as Reads Per Thousand, normalized to the reads mapped on the nuclear and mitochondrial genomes. The average of two biological replicates is presented \pm SD.

Region	sRNA-Seq					WTSS	
	T0	Linc 10m	T0	CAP 15min	CAP 3h	T0	Linc 10min
5'UTR (89 bp)	6.15 \pm 0.88	7.81 \pm 0.67	1.21 \pm 0.71	16.0 \pm 0.16	10.2 \pm 0.25	0.04 \pm 0.0	0.05 \pm 0.0
Exons (1057 bp)	0.33 \pm 0.002	0.28 \pm 0.004	0.23 \pm 0.002	0.54 \pm 0.001	1.60 \pm 0.41	188.9 \pm 16.9	208.4 \pm 2.6
Introns (5591 bp)	0.15 \pm 0.002	0.16 \pm 0.005	0.12 \pm 0.003	0.38 \pm 0.03	0.42 \pm 0.01	6.9 \pm 0.65	5.25 \pm 0.2
3'UTR (31 bp)	0.001 \pm 0	0.001 \pm 0	0.0002 \pm 0	0.001 \pm 0.001	0.004 \pm 0	0.58 \pm 0.26	1.0 \pm 0.18

Supplementary Table S8. Effect of transcription inhibition on sRNAs and RNA levels of *psbA*. sRNA and RNA levels are reported as the sum of the reads normalized to the total mapped of the nuclear and mitochondrial genomes for each feature and averaged between two biological replicates \pm S.D.

Region		sRNA-Seq (RPM)					
		Mixotrophic			Phototrophic		
Length (bp)	Type	T0	Rif 6h	log ₂ FC	T0	Rif 6h	log ₂ FC
89	5'UTR	6153 \pm 882.6	11369.9 \pm 960.4	0.9	4098.8 \pm 598.0	5968.4 \pm 2081.3	0.5
1057	Exons	325.7 \pm 2.4	75.4 \pm 0.1	-2.1	274.0 \pm 15.7	240.8 \pm 33.2	-0.2
5591	Introns	145.4 \pm 2.0	19.8 \pm 3.4	-2.9	131.8 \pm 0.2	13.8 \pm 1	-3.3
31	3'UTR	1.1 \pm 0.4	0.6 \pm 0.1	-0.9	1.0 \pm 0.3	1.1 \pm 0.1	0.1

Region		WTSS (RPM \times 10 ³)					
		Mixotrophic			Phototrophic		
Length (bp)	Type	T0	Rif 6h	FC	T0	Rif 6h	FC
89	5'UTR	0.04 \pm 0.0	0.004 \pm 0.2	-3.2	0.04 \pm 0.0	0.0 \pm 0.0	-5.7
1057	Exons	188.9 \pm 16.9	121.5 \pm 19.8	-0.6	243.8 \pm 7.4	142.8 \pm 15.2	-0.8
5591	Introns	6.9 \pm 0.6	1.4 \pm 0.2	-2.2	6.4 \pm 0.6	1.5 \pm 0.6	-2.0
31	3'UTR	0.58.5 \pm 0.2	0.5 \pm 0.2	-0.002	0.8 \pm 0.1	0.7 \pm 0.1	-0.2

Supplementary Table S9. Differential expression analysis: WTSS and sRNA-Seq coverage of protein-coding genes (sense strand).

(see “supplementary_data_1.xlsx”, attached)

Supplementary Table S10. RPM of all sRNA-Seq reads, RPFs and non-RPFs mapped to CDS of Cp protein coding genes after CAP and Rif treatment in WT strains.

(see “supplementary_data_1.xlsx”, attached)

Supplementary Table S11. Differential expression analysis for protein-coding genes: WTSS coverage between mixo- and photo-trophic conditions.

(see “supplementary_data_1.xlsx”, attached)

Supplementary Table S12. Mean expression levels of Cp genes in mixo- or photo-trophic conditions. Averaged RPKM values of mixo- and phototrophic WTSS datasets computed separately on the CDS, on the three exons of *psaA* and on the *WendyB* and *tscA* genes. The value for *psbA* is the average of the RPKM computed independently for the five exons. "Group" indicates the expression level category based on a RPKM arbitrary threshold. Differentially expressed genes between mixo- and phototrophic growth are marked with an asterisk.

WT t222	WT t222				
Gene	RPKM	Group	Gene	RPKM	Group
<i>psbA</i>	108939	High	<i>orf8</i>	1647	Moderate
<i>rbcL</i>	63132	High	<i>rpl2</i>	1779	Moderate
<i>atpH</i>	37560	High	<i>rps11*</i>	1326	Moderate
<i>psbD</i>	18708	High	<i>rps18</i>	1943	Moderate
<i>atpA</i>	13869	High	<i>rps7</i>	1548	Moderate
<i>psbC</i>	12894	High	<i>atpF</i>	1526	Moderate
<i>psbB</i>	12294	High	<i>rps2</i>	1770	Moderate
<i>psaA-3</i>	10603	High	<i>orf5</i>	1215	Moderate
<i>atpB</i>	11031	High	<i>rps8</i>	1155	Moderate
<i>psbT</i>	10876	High	<i>rps19</i>	1291	Moderate
<i>psbK</i>	7631	Moderate	<i>rpl14</i>	1012	Moderate
<i>petG</i>	8465	Moderate	<i>rpl16</i>	1001	Moderate
<i>psbE</i>	7952	Moderate	<i>tscA</i>	1391	Moderate
<i>psbL</i>	7529	Moderate	<i>rps9</i>	759	Low
<i>psaB</i>	6332	Moderate	<i>orf7</i>	986	Low
<i>psbJ</i>	5555	Moderate	<i>rps14</i>	871	Low
<i>psaC</i>	6120	Moderate	<i>ycf4</i>	730	Low
<i>psbM</i>	6440	Moderate	<i>ycf3</i>	737	Low
<i>petD</i>	6991	Moderate	<i>cemA</i>	761	Low
<i>rps12*</i>	4947	Moderate	<i>rps3</i>	741	Low
<i>petL</i>	6875	Moderate	<i>chlN</i>	620	Low
<i>psaA-2</i>	3896	Moderate	<i>clpP</i>	469	Low
<i>petA</i>	4499	Moderate	<i>chlL</i>	454	Low
<i>tufA*</i>	3458	Moderate	<i>rpl23</i>	525	Low
<i>ycf12</i>	6731	Moderate	<i>psbl</i>	701	Low
<i>psbZ</i>	3833	Moderate	<i>rpoB-1</i>	514	Low
<i>petB</i>	3691	Moderate	<i>WendyA*</i>	273	Low
<i>psbH</i>	3687	Moderate	<i>ycf1</i>	317	Low
<i>atpI</i>	3270	Moderate	<i>chlB</i>	278	Low
<i>I-Cre*</i>	2397	Moderate	<i>rpoA</i>	261	low
<i>rpl36</i>	3227	Moderate	<i>rpoB-2</i>	266	low
<i>atpE</i>	3602	Moderate	<i>rpoC-2</i>	194	low
<i>psbF</i>	2598	Moderate	<i>rpoC-1</i>	153	low
<i>psaA-1</i>	2549	Moderate	<i>rps4</i>	147	low
<i>psaJ</i>	3787	Moderate	<i>psbN</i>	153	low
<i>rpl20</i>	2996	Moderate	<i>orf528*</i>	98	low
<i>rpl5</i>	2030	Moderate	<i>ccsA</i>	117	low
<i>orf58</i>	2182	Moderate	<i>ycf2</i>	90	low
<i>orf6*</i>	2121	Moderate	<i>WendyB</i>	2	low

Supplementary Table S13. Differential expression analysis: WTSS coverage for nuclear genes.

(see “supplementary_data_2.xlsx”, attached)

Supplementary Table S14. Comparison of WTSS and sRNA-Seq coverage on sense and antisense strands for Cp regions.

(see “supplementary_data_3.xlsx”, attached)

Supplementary Table S15. Differential expression analysis: sRNA-Seq coverage of protein-coding genes (antisense strand).

(see “supplementary_data_2.xlsx”, attached)

Supplementary Table S16. sRNA-Seq results of CAP-treated WT cc-4533 and mutants in TCA1, PPR and OPR genes (sense strand).

(see “supplementary_data_3.xlsx”, attached)

Supplementary references

- Anthonisen, I.L., Salvador, M.L. and Klein, U. (2001) Specific sequence elements in the 5' untranslated regions of rbcL and atpB gene mRNAs stabilize transcripts in the chloroplast of *Chlamydomonas reinhardtii*. *Rna*, 7, 1024-1033.
- Blowers, A.D., Klein, U., Ellmore, G.S. and Bogorad, L. (1993) Functional in vivo analyses of the 3' flanking sequences of the *Chlamydomonas* chloroplast rbcL and psaB genes. *Mol.Gen.Genet.*, 238, 339-349.
- Boudreau, E., Takahashi, Y., Lemieux, C., Turmel, M. and Rochaix, J.D. (1997) The chloroplast ycf3 and ycf4 open reading frames of *Chlamydomonas reinhardtii* are required for the accumulation of the photosystem I complex. *EMBO J*, 16, 6095-6104.
- Bruick, R.K. and Mayfield, S.P. (1998) Processing of the psbA 5' untranslated region in *Chlamydomonas reinhardtii* depends upon factors mediating ribosome association. *The J Cell Biol*, 143, 1145-1153.
- Chen, Z. Y. and J. V. Moroney (1995). "Identification of a *Chlamydomonas reinhardtii* chloroplast gene with significant homology to bacterial genes involved in cytochrome c biosynthesis." *Plant Physiol* 108(2), 843-844.
- Cui, L., Leebens-Mack, J., Wang, L.S., Tang, J., Rymarquis, L., Stern, D.B. and dePamphilis, C.W. (2006) Adaptive evolution of chloroplast genome structure inferred using a parametric bootstrap approach. *BMC Evol Biol*, 6, 13.
- Drager, R.G., Girard-Bascou, J., Choquet, Y., Kindle, K.L. and Stern, D.B. (1998) In vivo evidence for 5'-->3' exoribonuclease degradation of an unstable chloroplast mRNA. *Plant J*, 13, 85-96.
- Drager, R.G., Higgs, D.C., Kindle, K.L. and Stern, D.B. (1999) 5' to 3' exoribonucleolytic activity is a normal component of chloroplast mRNA decay pathways. *Plant J*, 19, 521-531.
- Drapier, D., Girard-Bascou, J. and Wollman, F.A. (1992) Evidence for Nuclear Control of the Expression of the atpA and atpB Chloroplast Genes in *Chlamydomonas*. *Plant Cell*, 4, 283-295.
- Drapier, D., Suzuki, H., Levy, H., Rimbault, B., Kindle, K.L., Stern, D.B. and Wollman, F.A. (1998) The chloroplast atpA gene cluster in *Chlamydomonas reinhardtii*. Functional analysis of a polycistronic transcription unit. *Plant Physiol*, 117, 629-641.
- Erickson, J.M., Rahire, M. and J.D., R. (1984) *chlamydomonas reinhardtii* gene for the 32 000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. *EMBO J*, 3, 2753-2762.
- Erickson, J.M., Rahire, M., Malnoë, P., Girard-Bascou J., Pierre, Y., Bennoun, P., Rochaix, J.D. (1986) Lack of the D2 protein in a *Chlamydomonas reinhardtii* psbD mutant affects photosystem II stability and D1 expression. *EMBO J*. 5(8),1745-54.
- Fan, W.H., Woelfle, M.A. and Mosig, G. (1995) Two copies of a DNA element, 'Wendy', in the chloroplast chromosome of *Chlamydomonas reinhardtii* between rearranged gene clusters. *Plant Mol Biol*, 29, 63-80.
- Girard-Bascou, J., Y. Choquet, M. Schneider, M. Delosme and M. Dron (1987). Characterization of a chloroplast mutation in the psaA2 gene of *Chlamydomonas reinhardtii*. *Curr. Genet.* 12(7), 489-495.
- Glanz, S., Jacobs, J., Kock, V., Mishra, A. and Kück, U. (2012). Raa4 is a trans-splicing factor that specifically binds chloroplast tscA intron RNA. *Plant J* 69(3): 421-431.
- Goldschmidt-Clermont, M., Choquet, Y., Girard-Bascou, J., Michel, F., Schirmer-Rahire, M. and Rochaix, J.D. (1991) A small chloroplast RNA may be required for trans-splicing in *Chlamydomonas reinhardtii*. *Cell*, 65, 135-143.
- Goldschmidt-Clermont, M., Rahire, M. and Rochaix, J.D. (2008) Redundant cis-acting determinants of 3' processing and RNA stability in the chloroplast rbcL mRNA of *Chlamydomonas*. *Plant J*, 53, 566-577.
- Higgs, D.C., Shapiro, R.S., Kindle, K.L. and Stern, D.B. (1999) Small cis-acting sequences that specify secondary structures in a chloroplast mRNA are essential for RNA stability and translation. *Mol.Cell Biol.*, 19, 8479-8491.
- Higgs, D.C., Kuras, R., Kindle, K.L., Wollman, F.A., Stern, D.B. (1998) Inversions in the *Chlamydomonas* chloroplast genome suppress a petD 5' untranslated region deletion by creating functional chimeric mRNAs. *Plant J*, 14(6),663-71.
- Huang, C., Wang, S., Chen, L., Lemieux, C., Otis, C., Turmel, M. and Liu, X. Q. (1994) The *Chlamydomonas* chloroplast clpP gene contains translated large insertion sequences and is essential for cell growth. *Mol. Gen. Genet.* 244(2), 151-159.
- Jiao, H.S., Hicks, A., Simpson, C., Stern, D.B. (2004). Short dispersed repeats in the *Chlamydomonas* chloroplast genome are collocated with sites for mRNA 3' end formation. *Curr Genet.*, 45(5),311-22
- Johnson, C.H. and Schmidt, G.W. (1993) The psbB gene cluster of the *Chlamydomonas reinhardtii* chloroplast: sequence and transcriptional analyses of psbN and psbH. *Plant Molecular Biology*, 22, 645-658.

- Johnson, X., Wostrikoff, K., Finazzi, G., Kuras, R., Schwarz, C., Bujaldon, S., Nickelsen, J., Stern, D.B., Wollman, F.A. and Vallon, O. (2010) MRL1, a conserved Pentatricopeptide repeat protein, is required for stabilization of rbcL mRNA in Chlamydomonas and Arabidopsis. *Plant Cell*, 22, 234-248.
- Klein, U., De Camp, J.D. and Bogorad, L. (1992) Two types of chloroplast gene promoters in Chlamydomonas reinhardtii. *Proc.Natl.Acad.Sci.U.S.A.*, 89, 3453-3457
- Komine, Y., Kwong, L., Anguera, M.C., Schuster, G. and Stern, D.B. (2000) Polyadenylation of three classes of chloroplast RNA in Chlamydomonas reinhardtii. *Rna*, 6, 598-607.
- Lee, H., Bingham, S.E. and Webber, A.N. (1996) Function of 3' non-coding sequences and stop codon usage in expression of the chloroplast psaB gene in Chlamydomonas reinhardtii. *Plant Molecular Biology*, 31, 337-354.
- Li, J., Goldschmidt-Clermont, M. and Timko, M. P. (1993) Chloroplast-encoded chlB is required for light-independent protochlorophyllide reductase activity in Chlamydomonas reinhardtii. *Plant Cell* 5(12), 1817-1829.
- Liu, X. Q., Gillham, N. W. and Boynton, J. E. (1989) Chloroplast ribosomal protein gene rps12 of Chlamydomonas reinhardtii. Wild-type sequence, mutation to streptomycin resistance and dependence, and function in Escherichia coli. *J Biol Chem*. 27, 16100-16108.
- Loiselay, C., Gumpel, N.J., Girard-Bascou, J., Watson, A.T., Purton, S., Wollman, F.A. and Choquet, Y. (2008) Molecular identification and function of cis- and trans-acting determinants for petA transcript stability in Chlamydomonas reinhardtii chloroplasts. *Mol Cell Biol*, 28, 5529-5542.
- Loizeau, K., Qu, Y., Depp, S., Fiechter, V., Ruwe, H., Lefebvre-Legendre, L., Schmitz-Linneweber, C. and Goldschmidt-Clermont, M. (2014) Small RNAs reveal two target sites of the RNA-maturation factor Mbb1 in the chloroplast of Chlamydomonas. *Nucleic Acids Res*, 42, 3286-3297.
- Majeran, W., Olive, J., Drapier, D., Vallon, O. and Wollman, F.A. (2001) The light sensitivity of ATP synthase mutants of Chlamydomonas reinhardtii. *Plant Physiol*, 126, 421-433.
- Matsumoto, T., M. Matsuo, and Matsuda, Y. (1991). Structural analysis and expression during dark-light transitions of a gene for cytochrome f in Chlamydomonas reinhardtii. *Plant Cell Physiol*. 32863-872.
- Monod, C., Goldschmidt-Clermont, M. and Rochaix, J.D. (1992) Accumulation of chloroplast psbB RNA requires a nuclear factor in Chlamydomonas reinhardtii. *Mol.Gen.Genet.*, 231, 449-459.
- Murakami, S., Kuehnle, K. and Stern, D.B. (2005) A spontaneous tRNA suppressor of a mutation in the Chlamydomonas reinhardtii nuclear MCD1 gene required for stability of the chloroplast petD mRNA. *Nucleic Acids Res*, 33, 3372-3380.
- Nickelsen, J., Fleischmann, M., Boudreau, E., Rahire, M. and Rochaix, J.D. (1999) Identification of cis-acting RNA leader elements required for chloroplast psbD gene expression in Chlamydomonas. *Plant Cell*, 11, 957-970.
- Nickelsen, J., van Dillewijn, J., Rahire, M. and Rochaix, J.D. (1994) Determinants for stability of the chloroplast psbD RNA are located within its short leader region in Chlamydomonas reinhardtii. *Embo J*, 13, 3182-3191.
- Ossenbuhl, F., Hartmann, K. and Nickelsen, J. (2002) A chloroplast RNA binding protein from stromal thylakoid membranes specifically binds to the 5' untranslated region of the psbA mRNA. *Eur J Biochem*, 269, 3912-3919.
- Randolph-Anderson, B. L., Boynton, J. E., Gillham, N. W., Huang, C. and Liu, X. Q. (1995) The chloroplast gene encoding ribosomal protein S4 in Chlamydomonas reinhardtii spans an inverted repeat-unique sequence junction and can be mutated to suppress a streptomycin dependence mutation in ribosomal protein S12. *Mol Gen Genet* , 247(3), 295-305.
- Robertson, D., Boynton, J.E. and Gillham, N.W. (1990) Cotranscription of the wild-type chloroplast atpE gene encoding the CF1/CF0 epsilon subunit with the 3' half of the rps7 gene in Chlamydomonas reinhardtii and characterization of frameshift mutations in atpE. *Mol Gen Genet*, 221, 155-163.
- Rochaix, J.D., Kuchka, M., Mayfield, S., Schirmer-Rahire, M., Girard-Bascou, J. and Bennoun, P. (1989) Nuclear and chloroplast mutations affect the synthesis or stability of the chloroplast psbC gene product in Chlamydomonas reinhardtii. *Embo J*, 8, 1013-1021.
- Sattarzadeh, A., Fuller, J., Moguel, S., Wostrikoff, K., Sato, S., Covshoff, S., Clemente, T., Hanson, M. and Stern, D.B. (2010) Transgenic maize lines with cell-type specific expression of fluorescent proteins in plastids. *Plant Biotechnol J*, 8, 112-125.
- Schneider, M., Darlix, J.L., Erickson, J. and Rochaix, J.D. (1985) Sequence organization of repetitive elements in the flanking regions of the chloroplast ribosomal unit of Chlamydomonas reinhardtii. *Nucleic Acids Res.*, 13, 8531-8541.
- Silk, G.W. and Wu, M. (1993) 3' transcript mapping of the Chlamydomonas reinhardtii chloroplast psbK mRNA using a novel method to prepare 3' end-labeled single-stranded DNA probes. *Biotechniques*, 14, 830-833.

- Silk, G.W., Dela Cruz, F. and Wu, M. (1990) Nucleotide sequence of the chloroplast gene for the 4 kD K polypeptide of photosystem II (psbK) and the psbK-tufA intergenic region of Chlamydomonas reinhardtii. *Nucleic Acids Res*, 18, 4930.
- Stampacchia, O., Girard-Bascou, J., Zanasco, J.L., Zerges, W., Bennoun, P. and Rochaix, J.D. (1997) A nuclear-encoded function essential for translation of the chloroplast psaB mRNA in chlamydomonas. *Plant Cell*, 9, 773-782.
- Sturm, N.R., Kuras, R., Buschlen, S., Sakamoto, W., Kindle, K.L., Stern, D.B. and Wollman, F.A. (1994) The petD gene is transcribed by functionally redundant promoters in Chlamydomonas reinhardtii chloroplasts. *Mol Cell Biol*, 14, 6171-6179.
- Swiatek, M., Kuras, R., Sokolenko, A., Higgs, D., Olive, J., Cinque, G., Muller, B., Eichacker, L.A., Stern, D.B., Bassi, R. et al. (2001) The chloroplast gene ycf9 encodes a photosystem II (PSII) core subunit, PsbZ, that participates in PSII supramolecular architecture. *Plant Cell*, 13, 1347-1367.
- Takahashi, Y., Matsumoto, H., Goldschmidt-Clermont, M. and Rochaix, J.D. (1994) Directed disruption of the Chlamydomonas chloroplast psbK gene destabilizes the photosystem II reaction center complex. *Plant Mol Biol*, 24, 779-788.
- Turmel, M., Choquet, Y., Goldschmidt-Clermont, M., Rochaix, J.D., Otis, C. and Lemieux, C. (1995) The trans-spliced intron 1 in the psaA gene of the Chlamydomonas chloroplast: a comparative analysis. *Curr Genet*, 27, 270-279.
- Vaistij, F.E., Goldschmidt-Clermont, M., Wostrikoff, K. and Rochaix, J.D. (2000) Stability determinants in the chloroplast psbB/T/H mRNAs of Chlamydomonas reinhardtii. *Plant J*, 21(5), 469-82., 21, 469-482.
- Wang, F., Johnson, X., Cavaiuolo, M., Bohne, A.V., Nickelsen, J. and Vallon, O. (2015) Two Chlamydomonas OPR proteins stabilize chloroplast mRNAs encoding small subunits of photosystem II and cytochrome b6 f. *Plant J*, 82, 861-873.
- Woessner, J.P., Gillham, N.W. and Boynton, J.E. (1986) The sequence of the chloroplast atpB gene and its flanking regions in Chlamydomonas reinhardtii. *Gene*, 44, 17-28.
- Zicker, A.A., Kadakia, C.S. and Herrin, D.L. (2007) Distinct roles for the 5' and 3' untranslated regions in the degradation and accumulation of chloroplast tufA mRNA: identification of an early intermediate in the in vivo degradation pathway. *Plant Mol Biol*, 63, 689-702.
- Zito, F., Kuras, R., Choquet, Y., Kössel, H. and Wollman, F. A. (1997). Mutations of cytochrome b6 in Chlamydomonas reinhardtii disclose the functional significance for a proline to leucine conversion by petB editing in maize and tobacco. *Plant Mol Biol* 33(1), 79-86.
- Xu, R., Bingham, S.E., Webber, A.N. (1993). Increased mRNA accumulation in a psaB frame-shift mutant of Chlamydomonas reinhardtii suggests a role for translation in psaB mRNA stability. *Plant molecular biology* 22(3), 465-474.