

# RNA-Seq Analysis of Sulfur-Deprived *Chlamydomonas* Cells Reveals Aspects of Acclimation Critical for Cell Survival<sup>W</sup>

David González-Ballester,<sup>a,1</sup> David Casero,<sup>b</sup> Shawn Cokus,<sup>b</sup> Matteo Pellegrini,<sup>b,c</sup> Sabeeha S. Merchant,<sup>c,d</sup> and Arthur R. Grossman<sup>a</sup>

<sup>a</sup> Department of Plant Biology, Carnegie Institution for Science, Stanford, California 94305

<sup>b</sup> Department of Molecular, Cell, and Developmental Biology, University of California, Los Angeles, California 90095-1606

<sup>c</sup> Institute of Genomics and Proteomics, University of California, Los Angeles, California 90095

<sup>d</sup> Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1569

**The *Chlamydomonas reinhardtii* transcriptome was characterized from nutrient-replete and sulfur-depleted wild-type and *snrk2.1* mutant cells. This mutant is null for the regulatory Ser-Thr kinase SNRK2.1, which is required for acclimation of the alga to sulfur deprivation. The transcriptome analyses used microarray hybridization and RNA-seq technology. Quantitative RT-PCR evaluation of the results obtained by these techniques showed that RNA-seq reports a larger dynamic range of expression levels than do microarray hybridizations. Transcripts responsive to sulfur deprivation included those encoding proteins involved in sulfur acquisition and assimilation, synthesis of sulfur-containing metabolites, Cys degradation, and sulfur recycling. Furthermore, we noted potential modifications of cellular structures during sulfur deprivation, including the cell wall and complexes associated with the photosynthetic apparatus. Moreover, the data suggest that sulfur-deprived cells accumulate proteins with fewer sulfur-containing amino acids. Most of the sulfur deprivation responses are controlled by the SNRK2.1 protein kinase. The *snrk2.1* mutant exhibits a set of unique responses during both sulfur-replete and sulfur-depleted conditions that are not observed in wild-type cells; the inability of this mutant to acclimate to S deprivation probably leads to elevated levels of singlet oxygen and severe oxidative stress, which ultimately causes cell death. The transcriptome results for wild-type and mutant cells strongly suggest the occurrence of massive changes in cellular physiology and metabolism as cells become depleted for sulfur and reveal aspects of acclimation that are likely critical for cell survival.**

## INTRODUCTION

The dominant form of sulfur (S) in terrestrial and aquatic habitats is usually the sulfate anion (SO<sub>4</sub><sup>2-</sup>), the most oxidized form of S. Animals do not have the enzymatic machinery needed for reducing SO<sub>4</sub><sup>2-</sup> to sulfide (S<sup>2-</sup>), which is required to synthesize most S-containing compounds. Plants and microbes have specific transporters that efficiently import SO<sub>4</sub><sup>2-</sup> into cells, where it is activated and then reduced to S<sup>2-</sup> for incorporation into S-containing amino acids and other molecules, such as S-adenosylmethionine (SAM), GSH, FeS clusters, thionucleosides, sulfolipids, vitamins, and enzyme cofactors, such as CoA (CoA), molybdopterin, thiamine, or biotin. However, most organisms have a low capacity to store S and thus have developed diverse acclimation strategies that optimize S use and balance the rate of S metabolism with S availability. These processes allow organisms to remain viable even when S is severely limiting to growth.

The alga *Chlamydomonas reinhardtii* (*Chlamydomonas* throughout) has been used as a model organism for studying

responses of photosynthetic eukaryotes to S deprivation. S-responsive processes in this alga include rapid changes in cell size (Zhang et al., 2002), elevated production of hydrolytic extracellular enzymes (de Hostos et al., 1989; Takahashi et al., 2001), alterations in cell wall structure (Takahashi et al., 2001), changes in the activities and composition of the photosynthetic apparatus (Collier and Grossman, 1992; Wykoff et al., 1998; Zhang et al., 2004), elevated SO<sub>4</sub><sup>2-</sup> transport activity (Yildiz et al., 1994; Pootakham et al., 2010), and the synthesis of enzymes (plus the encoding transcripts) required for efficient S assimilation (Ravina et al., 1999, 2002; González-Ballester and Grossman, 2009). Moreover, in the last 5 years, S-depleted *Chlamydomonas* cells have been used for microarray-based RNA abundance studies (Zhang et al., 2004; Nguyen et al., 2008), determination of metabolite profiles (Bolling and Fiehn, 2005), and sustained production of H<sub>2</sub> (Ghirardi et al., 2007). Many studies of S-deprived plants have also been performed (Lewandowska and Sirko, 2008; Kopriva et al., 2009); some of these studies include determinations of transcript levels and metabolite profiles (Hirai et al., 2003; Maruyama-Nakashita et al., 2003; Nikiforova et al., 2003, 2005a, 2005b; Lunde et al., 2008).

A number of regulatory elements controlling S deprivation responses in *Chlamydomonas* have been identified. The sulfur acclimation gene *SAC1* encodes a polypeptide with homology to Na<sup>+</sup>/SO<sub>4</sub><sup>2-</sup> transporters (SLC13 family). However, the protein appears to function as a sensor that responds to extracellular SO<sub>4</sub><sup>2-</sup> levels

<sup>1</sup> Address correspondence to q62gobad@uco.es.

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Arthur R. Grossman (arthurg@stanford.edu).

<sup>W</sup> Online version contains Web-only data.

www.plantcell.org/cgi/doi/10.1105/tpc.109.071167

rather than as a transporter (Davies et al., 1996; Moseley and Grossman, 2009). The *sac1* mutant exhibits marked reduction in accumulation of many transcripts encoding proteins associated with S acclimation responses (Takahashi et al., 2001; Ravina et al., 2002; Zhang et al., 2004) and rapidly loses viability following exposure to S deprivation; the loss of viability is a consequence of the inability of the cells to decrease photosynthetic electron transport activity (Wykoff et al., 1998). However, transcription from some S-responsive genes does not show an absolute dependence on SAC1. For example, the *sac1* mutant still develops a considerable level of high affinity  $\text{SO}_4^{2-}$  transport activity (Davies et al., 1996) and accumulates transcripts encoding  $\text{SO}_4^{2-}$  transporters in response to S deprivation (González-Ballester et al., 2008).

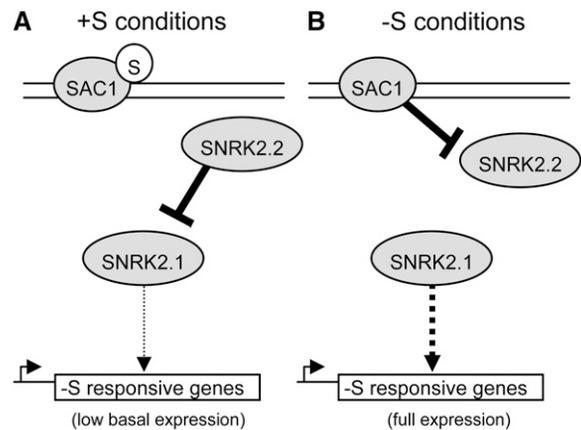
Two other genes encoding regulatory proteins that control S deprivation responses have been extensively characterized; they are the plant-specific SNF-1 related kinases *SNRK2.1* and *SNRK2.2* (the latter is also known as *SAC3*) (Davies et al., 1999; González-Ballester et al., 2008). *SNRK2.2* is responsible for repression of S-inducible genes when cells are replete for S (Davies et al., 1999; Ravina et al., 2002; González-Ballester et al., 2008). It may also be involved in repression of chloroplast transcription during S deprivation (Irihimovitch and Stern, 2006). The *SNRK2.1* kinase is required for most S-responsive gene expression and for maintaining cell viability during S deprivation. S-responsive gene expression and cell viability are more severely impacted in *snrk2.1* than in the *sac1* mutant during S deprivation (Pollock et al., 2005; González-Ballester et al., 2008). Moreover, while *SAC1* and *SNRK2.2* do not show a clear epistatic relationship, *SNRK2.1* is epistatic to *SNRK2.2*. Recently, it has been suggested that *SAC1* may act as a negative modulator of *SNRK2.2* activity during S deprivation (Moseley et al., 2009). Figure 1 shows a simplified diagram of the relationship among the different regulatory factors and the putative interactions associated with the acclimation of *Chlamydomonas* cells to S deprivation.

The recently developed RNA-seq technology (Wang et al., 2009) has made genome-wide transcript analyses both sensitive and quantitative. Here, we report on the use of RNA-seq technology to evaluate the responses of wild-type (the parental strain of the mutant) and the *snrk2.1* mutant of *Chlamydomonas* to S deprivation. A comparison of RNA-seq, microarray, and quantitative RT-PCR (qRT-PCR) results demonstrated reduced sensitivity of microarray analyses for estimating changes in mRNA abundance, especially for low abundance transcripts, relative to the RNA-seq platform. Furthermore, these RNA-seq studies have yielded large, quantitative data sets for transcript abundance in wild-type and *snrk2.1* mutant cells, both during S starvation and nutrient-replete growth. The data strongly suggest the occurrence of novel changes in cellular structures and metabolic pathways as wild-type cells acclimate to S deprivation. The data also define detailed responses of the *snrk2.1* mutant to S deprivation.

## RESULTS AND DISCUSSION

### Design, Statistical, and Comparative Analyses

We examined genome-wide transcript abundance in *Chlamydomonas* wild-type and *snrk2.1* mutant cells during growth in both



**Figure 1.** Model for S Deprivation-Responsive Gene Regulation.

(A) and (B) Simplified diagrams of gene regulation in wild-type cells in +S and -S conditions, respectively.

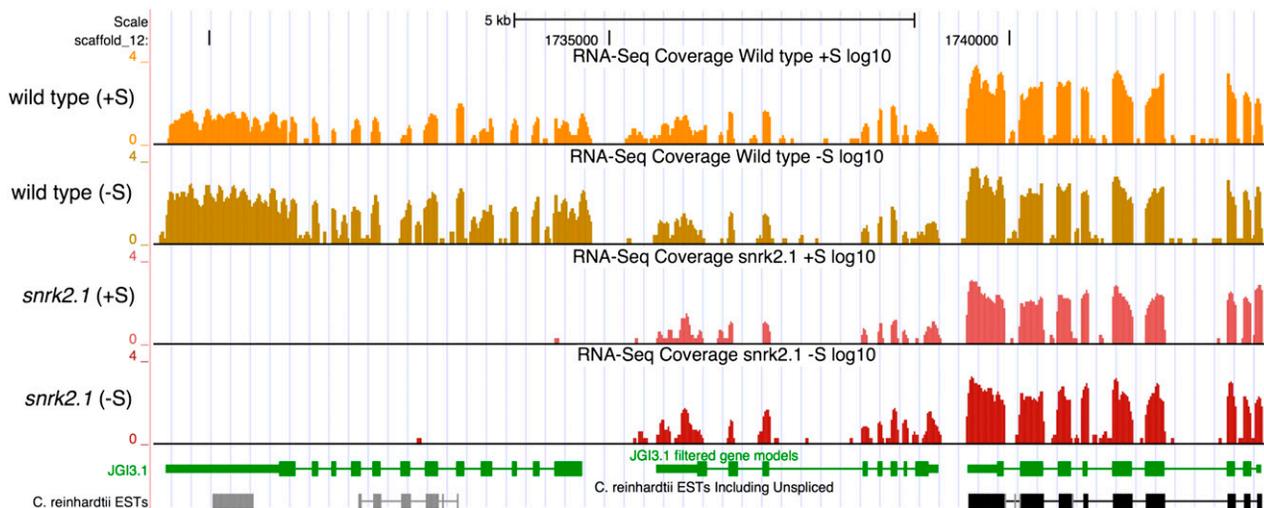
(A) In the presence of S, *SAC1* is inactive and *SNRK2.2* inhibits *SNRK2.1*-activated expression of S-responsive genes, leading to basal level expression.

(B) In S-deficient conditions, active *SAC1* blocks *SNRK2.2* inhibition of *SNRK2.1*, allowing full expression of S-responsive genes. This model is based on one previously developed by Moseley et al. (2009).

nutrient-replete (designated +S) and S-depleted (-S) conditions. RNA-seq and microarray profiling were used to determine the abundance of transcripts associated with each *Chlamydomonas* gene model (based on Joint Genome Institute [JGI] Chlre v.3.1) for each experimental condition (wt +S, wt -S, *snrk2.1* +S, and *snrk2.1* -S). The abundance of selected transcripts was also determined by qRT-PCR. Total RNA isolated from cells 6 h after they were transferred to -S and +S medium was used for RNA-seq. By 6 h after the cells were placed in -S conditions, the abundances of most of the induced transcripts had peaked. The same RNA preparations were also used for both microarrays and qRT-PCR analyses. Microarray hybridizations were also performed with RNA isolated from cells that were transferred from +S to both -S and +S conditions for 1 and 24 h. Genome-wide RNA-seq coverage graphs are available on a public installation of the UCSC genome browser at <http://genomes.mccdb.ucla.edu/CreSulfur/>. A browser image of a 14-kb region of the *Chlamydomonas* genome is shown in Figure 2.

Fold changes in transcript abundances using each of the techniques are presented in Supplemental Data Set 1 online. Transcripts with significant fold changes, as determined by RNA-seq, are grouped into categories, as shown in Figure 3 and Supplemental Data Set 2 online. An overview of the physiological functions related to these transcript categories is presented in Table 1, with an emphasis on differences between wild-type and *snrk2.1* mutant cells. Finally, some genes have been grouped according to their putative functionalities, independent of their relative expression levels (Supplemental Data Set 3 online).

Table 2 and Figure 4 show the correlations as fold change of transcript abundance for cells transferred from +S to -S conditions measured by RNA-seq, microarray hybridizations, and qRT-PCR. The levels of moderate and high abundance



**Figure 2.** RNA-Seq Coverage for a 14-kb Region of the *Chlamydomonas* Genome as Displayed by the UCSC Genome Browser.

Individual tracks for each *Chlamydomonas* strain and condition discussed in the article are displayed in different colors. The x axis represents a genome fragment with the linear arrangement of gene models along that fragment, while the y axis provides a representation of the degree of transcript coverage in log scale. JGI v.3.1 gene models and ESTs tracks are displayed at the bottom. See Supplemental Methods online for more details. The most 5' gene model (left-hand side) corresponds to the *AOT4* locus, a putative amino acid/polyamine transporter with increased expression in S-deprived wild-type cells that is barely expressed in *snrk2.1* cells (under both +S and –S conditions).

transcripts were similar for the RNA-seq and qRT-PCR results; there was significantly more variability in quantification of low abundance transcripts. For microarray values, the direction of change was the same compared with the other quantification methods, although the data exhibited a smaller dynamic range of fold change than RNA-seq or qRT-PCR (e.g., less of an increase at the upper range and much lower sensitivity at the lower range). Furthermore, low-abundance transcripts that could be quantified by RNA-seq and qRT-PCR were below the detection limits of the microarray. Finally, the most recent *Chlamydomonas* array queries <10,000 unique transcripts, while RNA-seq potentially covers the entire transcriptome. These results demonstrate that RNA-seq represents an excellent genome-scale platform for analyzing transcript levels, yielding fold changes that agree well with qRT-PCR quantifications. Unless otherwise noted, the results discussed in this manuscript will be from the +S and 6 h –S samples analyzed by RNA-seq, with many of the trends corroborated by qRT-PCR.

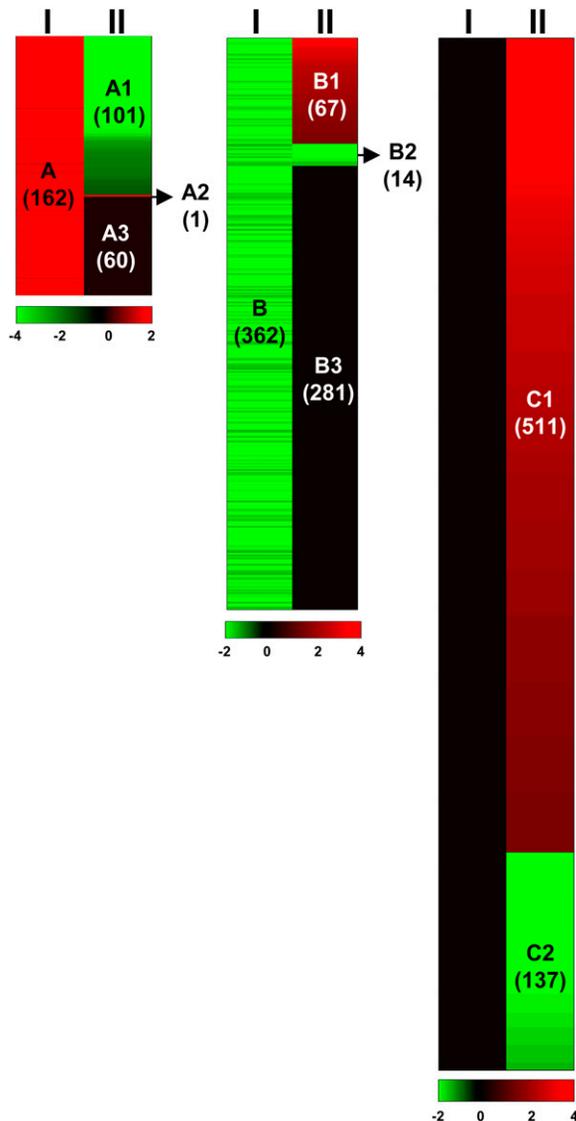
## Sulfur Metabolism

### Sulfur Acquisition and Assimilation

S deprivation of *Chlamydomonas* cells elicits an increase in the abundance of many transcripts encoding proteins involved in  $\text{SO}_4^{2-}$  acquisition and assimilation (Figure 5, Table 3). For example, levels of transcripts for a number of  $\text{SO}_4^{2-}$  transporters exhibited dramatic changes in response to S deprivation; transcripts encoding two putative  $\text{Na}^+/\text{SO}_4^{2-}$  transporters (SLT1 and SLT2) and one putative  $\text{H}^+/\text{SO}_4^{2-}$  transporter (SULTR2) markedly increased, while transcripts for three other putative  $\text{SO}_4^{2-}$  transporters (SLT3, SULTR1, and SULTR3) remained approximately

the same or declined during S deprivation; these results suggest that SLT1, SLT2, and SULTR2 are the main contributors to the inducible high-affinity  $\text{SO}_4^{2-}$  transport activity (Yildiz et al., 1994), a hypothesis that has recently been experimentally confirmed (Pootakham et al., 2010). In the *snrk2.1* mutant, there was no or little elevation in the levels of transcripts encoding proteins involved in  $\text{SO}_4^{2-}$  acquisition and assimilation, confirming previous results demonstrating that SNRK2.1 activity is crucial for regulating S acclimation responses (González-Ballester et al., 2008).

Moreover, there are transcripts encoding enzymes involved in S assimilation and the synthesis of S metabolites that show either little change in abundance or that decline under –S conditions. For example, transcripts encoding three isoforms of the O-acetylserine(thiol)-lyase (OASTL1, OASTL2, and OASTL3), APS kinase (APK), APS reductase (APR), sulfite reductase (SIR3), SAT2, and enzymes required for GSH synthesis and degradation did not significantly change under –S conditions (Figure 5, Table 3; see Supplemental Data Set 3 online). Interestingly, transcripts encoding proteins required for the Met/SAM cycle and for the biosynthesis of the S-containing cofactors thiamine and biotin decreased in abundance under –S conditions (Figure 5, Table 3). Similarly, the level of a transcript in *Arabidopsis thaliana* that likely functions in thiamine biosynthesis was diminished under –S conditions (Maruyama-Nakashita et al., 2003), although some transcripts encoding enzymes of the Met/SAM cycle increased in S-deprived *Arabidopsis* plants (Nikiforova et al., 2003). There have been no previous reports demonstrating transcriptional regulation of biotin biosynthesis in plants during S deprivation. Reduced levels of transcripts encoding enzymes required for Met biosynthesis in *Chlamydomonas* are congruent with the low levels of free Met present in S-deprived *Chlamydomonas* cells



**Figure 3.** Cluster Diagrams Depicting the Fold Change in Transcript Levels Determined by RNA-Seq.

The  $\log_2$  relative expression values of 1.5 and  $-1.5$  were selected as the thresholds to designate the categories of transcripts that accumulate (or are more abundant) and decline (or are less abundant), respectively, for wild type  $-S$ /wild type  $+S$  (I) and *snrk2.1*  $-S$ /wild type  $-S$  (II) conditions. For more details about each of the categories, see Methods. The number of different transcripts in each category is indicated in the appropriate sector. Areas are proportional to the numbers of genes within the specific category. A short description of each category is given below. A, transcripts that accumulate in wild-type cells in  $-S$  conditions; A1, transcripts that accumulate in wild-type cells in  $-S$  conditions and that are relatively less abundant in the mutant in  $-S$  conditions; A2, transcripts that accumulate in wild-type cells in  $-S$  conditions and that are relatively more abundant in the mutant in  $-S$  conditions; A3, transcripts that accumulate in wild-type cells to approximately the same extent as in mutant cells in  $-S$  conditions; B, transcripts that decline in wild-type cells in  $-S$  conditions; B1, transcripts that decline in wild-type cells in  $-S$  conditions and that are relatively more abundant in mutant cells in  $-S$  conditions; B2, transcripts that decline in wild-type cells in  $-S$  conditions

(Bolling and Fiehn, 2005), while Met levels in *Arabidopsis* plants subjected to S deprivation do not decline (Nikiforova et al., 2005b).

Like the transcriptional regulator SLIM1 from *Arabidopsis* (Maruyama-Nakashita et al., 2006), the *Chlamydomonas* SNRK2.1 kinase controls expression of many genes encoding proteins critical for  $SO_4^{2-}$  uptake, reductive assimilation of  $SO_4^{2-}$  to Cys, and S recycling (see section below: Redistribution and Recycling of S). However, transcriptional regulation of thiamine biosynthesis and Met/SAM cycle genes does not appear to be directly controlled by SNRK2.1, and there is no evidence suggesting that SLIM1 controls these genes in *Arabidopsis* (Maruyama-Nakashita et al., 2006); the activity of these genes may be impacted by the levels of S metabolites in both *Chlamydomonas* and *Arabidopsis*.

Despite the finding that Met levels in *Arabidopsis* do not decrease during S deprivation, SAM levels in *Arabidopsis* do decline (Nikiforova et al., 2005b); we would also expect the SAM levels in *Chlamydomonas* to decline, although they were not directly measured. SAM potentially represents a large drain on the Met pool; 80% of the Met synthesized in *Lemna paucicostata* cells may be consumed in the SAM pathway, with only  $\sim 20\%$  routed to protein synthesis (Giovanelli et al., 1985). Hence, reduced levels of SAM may impact several metabolic pathways including those responsible for the synthesis of thiamine, chlorophyll, polyamines, and lipids. Therefore, not unexpectedly and as previously observed in plants (Hirai et al., 2003; Maruyama-Nakashita et al., 2003; Nikiforova et al., 2003), the levels of numerous transcripts encoding proteins associated with these biosynthetic pathways in *Chlamydomonas* were significantly affected during S deprivation (see Supplemental Data Set 2 online).

In this study, we also identified two transcripts encoding putative amino acid/polyamine transporters (AOT2 and AOT4) that increased in S-deprived wild-type cells. The absolute abundance of the AOT4 transcript was  $\sim 8.5$  times higher than that of AOT2. Interestingly, the AOT4 transcript was almost undetectable in *snrk2.1* cells maintained in either  $+S$  or  $-S$  conditions (Figure 2, Table 3), indicating that this gene is under stringent SNRK2.1 control. An increase in the AOT4 transcript level did not occur in phosphorous-depleted ( $-P$ ) or nitrogen-depleted ( $-N$ ) cells (Figure 6A). Although the only amino acid-specific transport activity that has been demonstrated in *Chlamydomonas* is for Arg (Kirk and Kirk, 1978), the existence of additional amino acid transporters has been proposed (Perez-Alegre and Franco, 1998), and recently we demonstrated that wild-type *Chlamydomonas* cells, unlike the *snrk2.1* mutant, were able to grow slowly

and that are relatively less abundant in mutant cells in  $-S$  conditions; B3, transcripts that decline in wild-type cells and decline to approximately the same extent in mutant cells in  $-S$  conditions; C1, transcripts that are more abundant in mutant cells than in wild-type cells in  $-S$  conditions and that do not change significantly in wild-type cells in  $-S$  conditions; C2, transcripts that are less abundant in mutant cells than in wild-type cells in  $-S$  conditions and that do not change significantly in wild-type cells in  $-S$  conditions. Additional information for all genes included in this figure can be found in Table 1 and Supplemental Data Set 2 online.

**Table 1.** General Overview of Metabolic Pathways and Cellular Processes for Which Related Transcripts Accumulate or Decline in Wild-Type and *snrk2.1* Mutant Cells during –S Conditions

Specific in the Wild Type	Common in the Wild Type and <i>snrk2.1</i> Mutant	Specific in <i>snrk2.1</i> Mutant
<b>Upregulated</b>		
<b>A1</b>	<b>A2+A3</b>	<b>C1</b>
S acquisition and assimilation (to Cys) Routing and recycling of S Remodeling of PSII antenna (LHCBM9) Cell wall and periplasmic proteins (HAPs, ARSs, and ECPs) SQDG synthesis Oxidative pentose phosphate pathway <i>CAH9</i> Others	Stress-related LHC proteins Others	Oxidative stress (singlet oxygen) Proteolysis-related genes Vacuolar and vesicular transport Apoptosis Lipid degradation Pro degradation Putrescine biosynthesis Low CO <sub>2</sub> -inducible genes <sup>a</sup> Shikimate biosynthesis ( <i>SHKG1</i> ) CoA biosynthesis Others
<b>Downregulated</b>		
<b>B1</b>	<b>B2+B3</b>	<b>C2</b>
tRNA biosynthesis <sup>b</sup> Transcription and translation processes <sup>b</sup> Polyamine biosynthesis Others	Met, SAM, thiamine, and biotin biosynthesis Photosynthesis and chlorophyll biosynthesis Carbon and folate metabolism Plastid ribosomes <sup>c</sup> Amino acid biosynthesis <sup>c</sup> Lipid metabolism Purine and pyrimidine metabolism Others	S assimilation ( <i>APK</i> and <i>APR</i> ) Urea/Arg cycle Carbon fixation Others

A1, A2, A3, B1, B2, B3, C1, and C2 designate the same group categories described in Figure 3 and Supplemental Data Set 2 online. Footnotes indicate exceptions.

<sup>a</sup>Some transcripts were more abundant in *snrk2.1* mutant relative to wild-type cells in +S conditions.

<sup>b</sup>Some transcripts were also diminished in *snrk2.1* mutant cells.

<sup>c</sup>Some transcripts were not diminished in *snrk2.1* mutant cells.

with Cys, but not with Met, taurine, or homocysteine, as a sole S source (see Supplemental Figure 1 online). These results raise the possibility that the AOT2 and AOT4 transporters function in the uptake of Cys from the environment. Alternatively, they may participate in the redistribution of internal S-containing amino acids and/or polyamines among the different cellular compartments.

### Redistribution and Recycling of S

Several proteins encoded by transcripts that accumulated during –S conditions may be related to Cys degradation and/or to the recycling and redistribution of intracellular S (Figure 7). Cys dioxygenase (CDO) catalyzes the conversion of Cys to 3-sulfinoalanine (Figure 7), a key metabolite implicated in the control of Cys homeostasis and taurine biosynthesis in animal cells. The transcript from a putative *CDO1* gene increased in *Chlamydomonas* wild-type cells by ~28-fold during –S conditions (Table 3). Despite the potential importance of recycling S-containing amino acids, there are no reports to our knowledge that demonstrate CDO activity in plants, and the *CDO* gene has only recently been identified in bacteria (Dominy et al., 2006). Interestingly, the *CDO1* gene of *Chlamydomonas* is clustered in a tail-to-tail orientation with the *PWR1* gene (Pro-Trp-Arg motif) of unknown function whose transcript increased in abundance by ~76-fold during –S conditions. The levels of transcripts for both *CDO1* and *PWR1* are under the control of the SNRK2.1 kinase.

Taurine/α-ketoglutarate dioxygenase (TAUD), which catalyzes the conversion of taurine to SO<sub>3</sub><sup>2-</sup> and aminoacetaldehyde, may also be integral to the Cys degradation pathway (Figure 7). Two putative TAUD genes (*TAUD1* and *TAUD2*) were found on the *Chlamydomonas* genome; the transcripts from these genes increased by 142- and 37-fold, respectively, in S-deprived wild-type cells (Table 3). Exposure of the cells to –P and –N conditions did not alter *TAUD1* or *TAUD2* transcript levels (Figure 6B). The *TAUD1* and *TAUD2* genes are clustered together with another gene of unknown function (ID: 143415) that is also responsive to S levels; S deprivation elicited a 128-fold increase in the transcript from the latter gene. Furthermore, all three of the transcripts are under SNRK2.1 kinase control. While no TAUD homologs have been identified in plants or animals, the putative TAUD proteins of *Chlamydomonas* have significant similarity to bacterial and fungal TAUD/TfdA family proteins (~19 to 35% identity), as shown in Supplemental Figure 2 online. Moreover, the *Escherichia coli* TauD, the *Saccharomyces cerevisiae* YLL057c, and the *Pseudomonas syringae* AtsK α-ketoglutarate dioxygenases can efficiently metabolize various sulfonates in addition to taurine (Eichhorn et al., 1997; Hogan et al., 1999; Kahnert et al., 2000). Similarly, the putative TAUD proteins of *Chlamydomonas* might use alkanesulfonates other than taurine. For example, it is possible that TAUD functions in degrading chloroplast sulfolipids such as sulfoquinovosyl diacylglycerol (SQDG) (Figure 7); much of the sulfolipid may represent an S

**Table 2.** Comparative Quantification of Transcript Levels

ID <sup>a</sup>	Gene <sup>b</sup>	Wild Type –S/Wild Type +S			<i>snrk2.1</i> /Wild Type (+S)			<i>snrk2.1</i> /Wild Type (–S)		
		M	S	Q	M	S	Q	M	S	Q
205496	<i>ARS1</i>	3.6	10.5	9.5	nd	–1.2	–0.5	–3.0	–16.1	–12.0
205502	<i>SLT1</i>	2.6	6.9	7.9	nd	–6.3	–7.3	–5.1	–15.0	–20.4
205501	<i>SLT2</i>	4.2	6.3	9.3	nd	–2.5	–1.0	–4.4	–9.8	–10.3
150514	<i>SULTR2</i>	2.5	5.3	7.7	–0.9	–4.5	–3.9	–1.2	–6.4	–7.5
196910	<i>ATS1</i>	1.9	1.6	1.6	nd	–1.2	–1.3	–1.3	–3.0	–2.9
133924	<i>ATS2</i>	1.1	2.8	3.6	nd	0.0	0.0	–1.7	–4.4	–3.6
189320	<i>OASTL4</i>	2.0	2.9	5.6	–0.5	–0.3	–0.3	–1.6	–5.0	–6.6
205985	<i>SAT1</i>	3.9	5.2	4.2	nd	–0.7	0.1	–2.7	–4.2	–4.3
143892	<i>CDO1</i>	2.5	4.7	4.5	nd	–1.4	–4.0	nd	–7.3	–8.4
127464	<i>TAUD1</i>	3.3	7.1	5.6	nd	–3.5	–5.3	–1.3	–11.7	–12.5
184479	<i>LHCBM9</i>	2.9	10.0	8.8	nd	–2.6	–10.1	–3.8	–12.8	–20.5
130684	<i>ECP76</i>	3.0	10.7	10.3	nd	–0.8	–0.4	–2.5	–13.1	–11.8
143696	<i>HAP2</i>	5.2	8.4	7.1	nd	–0.1	–0.6	–5.6	–11.6	–11.5
183511	<i>RDP3</i>	2.8	4.1	4.4	nd	–0.8	–2.7	–0.8	–4.0	–6.6
182845	<i>SBDP</i>	2.1	5.4	5.0	–0.5	0.1	–1.6	–2.6	–8.3	–8.4
206105	<i>AOT4</i>	1.3	3.9	4.0	–2.6	–8.0	–8.8	–2.1	–13.0	–13.1

Fold change ( $\log_2$  values) of selected transcripts under specified conditions as determined by microarray (M), RNA-seq (S), and qRT-PCR (Q). Values are from cells maintained in nutrient-replete medium and exposed to –S conditions for 6 h. nd, not enough high-quality data for statistical tests.

<sup>a</sup>JGI protein accession number.

<sup>b</sup>The protein name associated with each of the genes is given in Table 3.

storage compound that can be rapidly mobilized when cells become S limited (Sugimoto et al., 2007). However, neither CDO nor TAUD polypeptides of *Chlamydomonas* appear to have presequences, suggesting that they function in the cytosol.

The *Chlamydomonas* genome also contains a number of other genes encoding putative sulfurtransferases that may be involved in the turnover of S compounds in the cell or in synthesizing new S-containing compounds. The rhodanases are sulfurtransferases thought to be involved in the synthesis of a diverse set of compounds, including molybdopterin, thiamin, CoA, biotin, lipoic acid, and iron-sulfur clusters (Kessler, 2006). *Chlamydomonas* has three genes (*RDP1*, 2, and 3) encoding small polypeptides that are part of a relatively poorly characterized family of rhodanase proteins (Bordo and Bork, 2002). The *RDP3* transcript rapidly and specifically accumulates when cells are deprived of S (~17-fold increase); this response is SNRK2.1 dependent (Table 3, Figure 6C). Moreover, Cys degradation can also occur via a 3-sulfinoalanine-independent pathway that generates a 3-mercaptopyruvate intermediate, which can be reduced by a mercaptopyruvate sulfotransferase (MST; designated TST in *Chlamydomonas*) (Figure 7). MST activity may also be part of the mechanism involved in detoxification of H<sub>2</sub>S in *Arabidopsis* (Riemenschneider et al., 2005). The MSTs/TSTs are related to rhodanases, with common sequence motifs and similar physicochemical properties (Colnaghi et al., 2001; Bordo and Bork, 2002), raising the possibility that *RDP3* is involved in the 3-sulfinoalanine-independent metabolism (Figure 7).

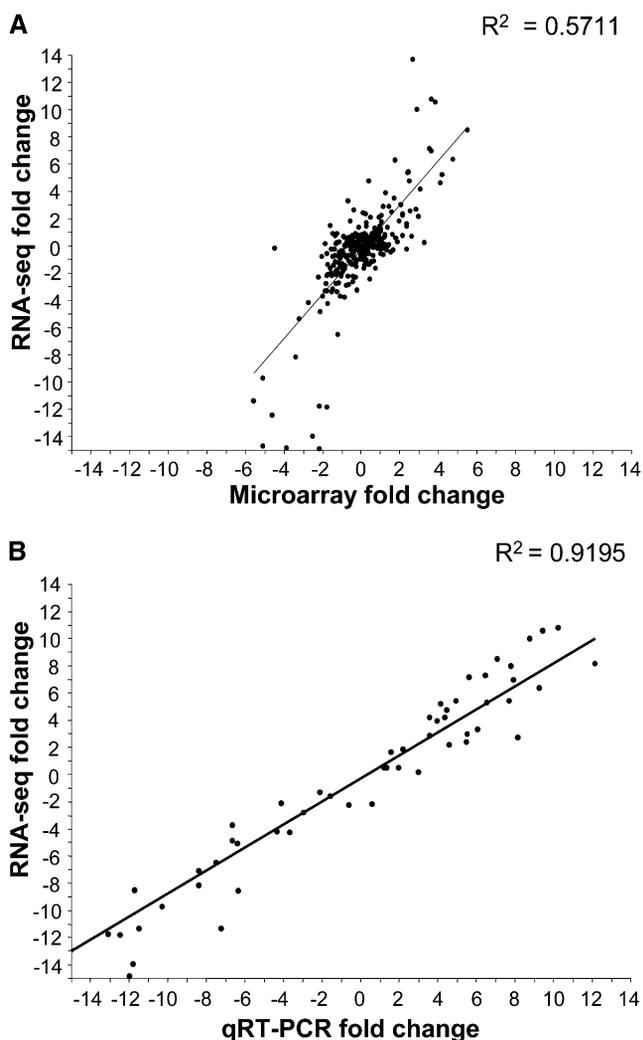
Cys desulfurases (CSDs), like the rhodanases, have also been proposed to act as sulfurtransferases. Like *RDP3*, the *Chlamydomonas* *CSD2* transcript increased when cells were deprived of S; this increase was SNRK2.1 dependent (Table 2). Hence, *RDP3* and *CSD2* may both be associated with redistribution of S

among pools of metabolites, which could be crucial under conditions in which cells are limited for S.

Alternatively, rhodanases and CSDs, which are able to bind selenium, may function in the detoxification of selenate or as selenium delivery proteins for the synthesis of selenocysteine, as has been suggested by others (Ogasawara et al., 2005; Mueller, 2006). The link between selenium and S homeostasis is supported by the findings that transcripts encoding several selenoproteins, including SELW1, MSP1, and two selenium binding proteins, accumulated in *Chlamydomonas* cells deprived of S and that this accumulation was dependent on SNRK2.1 (see Supplemental Data Set 2A online).

A major S-containing metabolite generated by pathways that recycle S from Cys is SO<sub>3</sub><sup>2-</sup>, which at high concentrations is cytotoxic. However, SO<sub>3</sub><sup>2-</sup> can be detoxified by oxidation to SO<sub>4</sub><sup>2-</sup> by sulfite oxidase (SUOX), as has been suggested for *Arabidopsis* SUOX (Hansch et al., 2007). *Chlamydomonas* has a *SUOX* gene that encodes an animal-type enzyme with a putative mitochondrial targeting sequence. Increased accumulation of the *SUOX* transcript during –S conditions supports the hypothesis that S deprivation in *Chlamydomonas* promotes the degradation of Cys and other S-containing metabolites. The SO<sub>4</sub><sup>2-</sup> generated by SUOX could be routed back into chloroplasts where it can be either incorporated into Cys by reductive assimilation or directly used for sulfation of molecules through the PAPS pathway. SUOX activity allows for SO<sub>4</sub><sup>2-</sup> regeneration in S-depleted cells and therefore may be required to maintain PAPS-dependent metabolism during –S conditions (Figure 7).

Although the substrate specificities of CDO1, TAUD1, TAUD2, RDP3, CSD2, and SUOX1 need to be clarified, the coordinated increases in their transcript levels in wild-type cells experiencing S deprivation indicates the critical nature of this group of



**Figure 4.** A Comparison of Fold Change ( $\log_2$ ) Values of Transcript Levels Measured by RNA-Seq, Microarrays, and qRT-PCR.

**(A)** Fold change ( $\log_2$ ) values for individual transcripts between two experimental conditions in which the RNA levels were quantified by microarray analyses ( $x$  axis) and were then compared with corresponding values calculated from RNA-seq experiments ( $y$  axis).

**(B)** Fold change ( $\log_2$ ) values for individual transcripts between two experimental conditions in which the RNA levels were quantified by qRT-PCR analyses ( $x$  axis) and then compared with corresponding values obtained by RNA-seq experiments ( $y$  axis). The pairs of experimental conditions compared include wild type  $-S$ /wild type  $+S$ ; *snrk2.1*  $-S$ /wild type  $-S$ ; *snrk2.1*  $+S$ /wild type  $+S$ , with all time points for  $-S$  corresponding to 6 h after the onset of S deprivation.

enzymes for the catabolism of Cys and/or sulfolipids, and redistributing intracellular S as the cells achieve a new homeostasis. Furthermore, the redistribution and recycling of S-containing molecules is likely under SNRK2.1 control since *CDO1*, *TAUD1*, *TAUD2*, *SUOX1*, *RDP3*, and *CSD2* transcripts are significantly diminished in the *snrk2.1* mutant relative to wild-type cells (Table 3).

### Sulfolipid Degradation and Synthesis

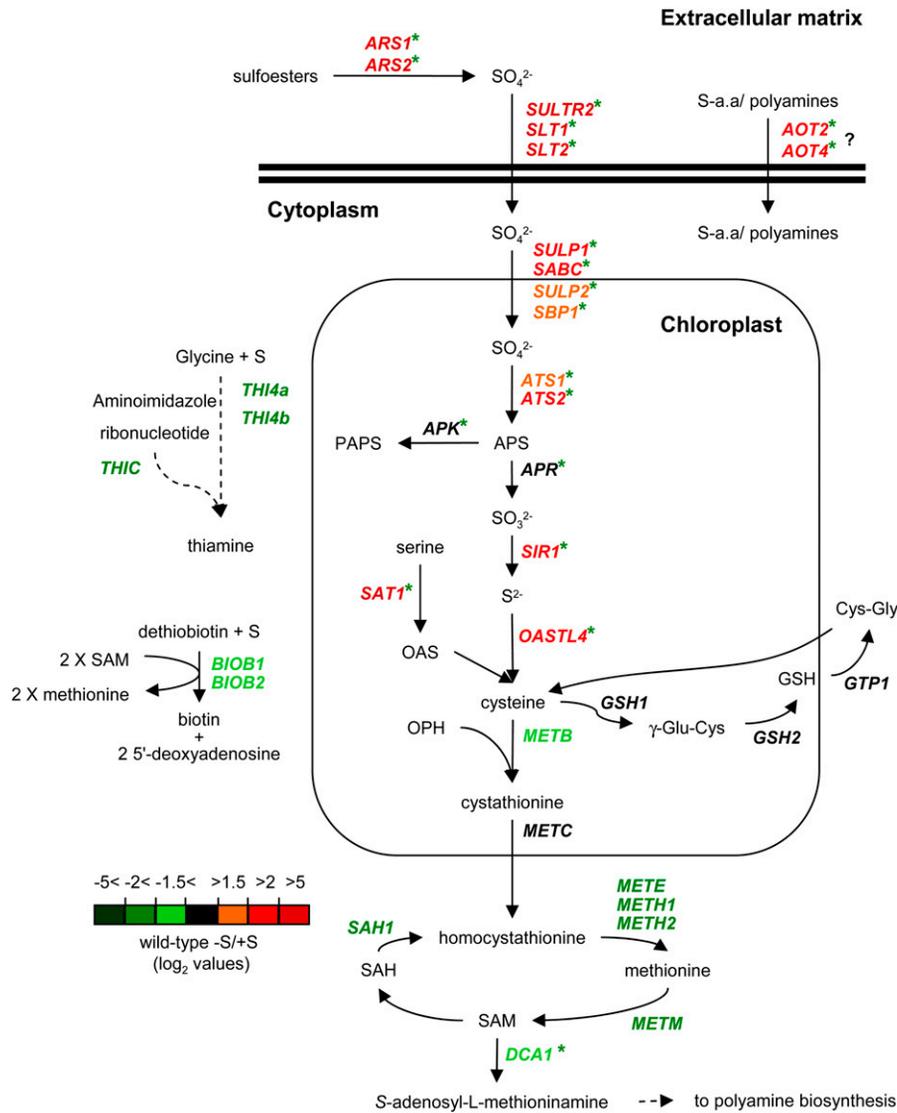
Whereas there is a rapid decline in the SQDG level in S-deprived *Chlamydomonas* cells (Sugimoto et al., 2007), the enzymes that catalyze SQDG degradation have not been identified, although they appear to be synthesized de novo after the cells are deprived of S (Sugimoto et al., 2009). The genes involved in this process may be under SNRK2.1 control and among those placed in category A1 in Figure 3 and in Supplemental Data Set 2 online. Transcripts potentially involved in sulfolipid degradation during exposure of *Chlamydomonas* to  $-S$  conditions encode TAUD1, TAUD2 (discussed above), and a putative lipase (JGI ID: 144554) (Table 3), although this possibility will require experimental validation. *Arabidopsis* plants also show some increase in the catabolism of sulfolipids during S deprivation. However, this degradation may reflect a general increase in lipid catabolism rather than a specific mechanism for recycling S (Nikiforova et al., 2005b); indeed, the SQDG levels in *Arabidopsis* plants transferred to  $-S$  conditions do not significantly decline (Okazaki et al., 2009). Instead, in *Arabidopsis* and other members of the Brassicaceae, glucosinolates may serve as an intracellular S reserve that can be rapidly catabolized under conditions of S limitation (Maruyama-Nakashita et al., 2003; Nikiforova et al., 2005b).

Surprisingly, S deprivation elicited an increase in three transcripts that have been associated with SQDG synthesis: *SQD1*, *SQD2*, and *LPB1* (Figure 7, Table 3). Increases in these transcripts were not observed in the *snrk2.1* mutant (this work); an increase in *SQD1* mRNA was also demonstrated to be impacted by both the *sac1* and *sac3* lesions (Zhang et al., 2004; Sugimoto et al., 2009). *SQD1* and *SQD2* encode UDP-sulfoquinovose synthase and sulfolipid synthase, respectively (Sato et al., 1995). *LPB1* (for Low Phosphate Bleaching protein) (Chang et al., 2005) is similar to UDP-glucose pyrophosphorylase 3 of *Arabidopsis*, which is involved in the synthesis of UDP-glucose, the precursor of sulfolipids (Okazaki et al., 2009). The *LPB1* transcript was elevated in cells exposed to  $-S$ , but not in cells exposed to  $-P$  or  $-N$  conditions (Figure 6D). Elevated levels of *SQD1*, *SQD2*, and *LPB1* transcripts during  $-S$  conditions are congruent with the increased capacity for SQDG biosynthesis previously observed in S-deprived *Chlamydomonas* cells (Sugimoto et al., 2009).

### Photosynthesis

S-deprived *Chlamydomonas* cells exhibit a marked decline in photosynthetic electron transport that has been associated with a SAC1-dependent decrease in photosystem II (PSII) activity (Wykoff et al., 1998). Moreover, *sac1* and *snrk2.1* mutants photobleach and die after  $\sim 3$  to 6 d of being transferred to medium devoid of S, while wild-type cells can survive for significantly longer periods (Davies et al., 1996; González-Ballester et al., 2008). These results suggest that both SAC1 and SNRK2.1 are important for tuning photosynthetic activity when the cells are starved for S and that this tuning is critical for cell viability.

Many transcripts encoding proteins associated with PSII, photosystem I (PSI), light-harvesting complexes (LHC) of PSI (LHCA) and PSII (LHCB), subunits of the cytochrome  $b_6f$  complex, photosynthetic electron transport, and chlorophyll biosynthesis declined when *Chlamydomonas* was exposed to  $-S$



**Figure 5.** Pathways for S Acquisition and Assimilation.

Genes encoding the proteins of these pathways are labeled with colors, indicating the fold change ( $\log_2$ ) of their transcript levels under S deprivation relative to nutrient-replete conditions for wild-type cells based on RNA-seq data (color code is given in the figure). The proteins encoded by each of the genes represented in the pathway diagram are given in Table 3. Dashed lines represent multiple metabolic steps. Asterisks represent genes that showed altered expression in *snrk2.1* mutant relative to wild-type cells upon S deprivation. APS, adenosine 5'-phosphosulfate; PAPS, adenosine 3'-phosphate 5'-phosphosulfate;  $S^{2-}$ , sulfide; OAS, O-acetyl-serine; OPH, O-phosphohomoserine; SAH, S-adenosyl-homocysteine; S-a.a, sulfur-containing amino acids.

conditions (see Supplemental Data Sets 2B and 3 online). A similar decrease in chlorophyll content and transcripts encoding proteins of the photosynthetic apparatus also occur in plants. Among the transcripts most sensitive to S deprivation were those encoding highly conserved members of the LHCBM protein family (Elrad and Grossman, 2004) that comprise the bulk of the peripheral PSII antennae. The mRNAs encoding LHCBM1-8 polypeptides declined after 6 h of S deprivation, with a more severe decline after 24 h (Figure 8A, Table 3). Levels of these transcripts were even lower in the *snrk2.1* mutant than in wild-type cells (Figure 8B, Table 3). By contrast, the LHCBM9 tran-

script, which was barely detectable when *Chlamydomonas* cells were grown in +S medium, increased by >1000-fold during S deprivation (see Supplemental Data Set 2 online) and was the second most abundant mRNA in wild-type S-depleted cells (see Supplemental Data Set 1 online). Recently, an increase in the *LHCBM9* transcript level during S deprivation was noted by others (Nguyen et al., 2008). This -S-dependent increase was not observed in the *snrk2.1* or *sac1* mutants (Figure 8C), suggesting that *LHCBM9* is a direct target of the S-sensing regulatory pathway; there was no increase in *LHCBM9* mRNA when wild-type cells were exposed to -P or -N conditions (Figure 6E).

**Table 3.** Expression Levels of Selected S-Regulated Genes

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type -S/ Wild Type +S		<i>snrk2.1</i> /Wild Type (+S)		<i>snrk2.1</i> /Wild Type (-S)		Wild Type	Wild Type +S	<i>snrk2.1</i> -S	<i>snrk2.1</i> +S
		S	Q	S	Q	S	Q	-S	+S	-S	+S
Sulfur acquisition and assimilation											
205496	<i>ARS1</i> , arylsulfatase (ext)	<b>10.5</b>	<b>9.5</b>	-1.2	-0.5	<b>-16.1</b>	<b>-12.0</b>	2717	2	0	1
55757	<i>ARS2</i> , arylsulfatase (ext)	<b>8.2</b>	<b>12.2</b>	1.0	<b>-2.9</b>	<b>-11.1</b>	<b>-15.1</b>	84	0	0	1
205502	<i>SLT1</i> , Na <sup>+</sup> /SO <sub>4</sub> <sup>2-</sup> transporter type (p)	<b>6.9</b>	<b>7.9</b>	<b>-6.3</b>	<b>-7.3</b>	<b>-15.0</b>	<b>-20.4</b>	2341	20	0	0
205501	<i>SLT2</i> , Na <sup>+</sup> /SO <sub>4</sub> <sup>2-</sup> transporter type (p)	<b>6.3</b>	<b>9.3</b>	<b>-2.5</b>	-1.0	<b>-9.8</b>	<b>-10.3</b>	1339	17	1	3
150514	<i>SULTR2</i> , H <sup>+</sup> /SO <sub>4</sub> <sup>2-</sup> transporter type (p)	<b>5.3</b>	<b>7.7</b>	<b>-4.5</b>	<b>-3.9</b>	<b>-6.4</b>	<b>-7.5</b>	152	4	2	0
196910	<i>ATS1</i> , ATP sulfurylase	<b>1.6</b>	<b>1.6</b>	-1.2	-1.3	<b>-3.0</b>	<b>-2.9</b>	2137	723	275	305
133924	<i>ATS2</i> , ATP sulfurylase (chl)	<b>2.8</b>	<b>3.6</b>	0.0	0.0	<b>-4.4</b>	<b>-3.6</b>	254	36	12	37
128906	<i>SULP1</i> , chl. SO <sub>4</sub> <sup>2-</sup> transport system	<b>2.4</b>		-0.6		<b>-3.4</b>		147	28	14	18
116547	<i>SULP2</i> , chl. SO <sub>4</sub> <sup>2-</sup> transport system	1.5		-0.7		<b>-2.2</b>		56	20	12	12
182359	<i>SBP1</i> , chl. SO <sub>4</sub> <sup>2-</sup> transport system	1.5		-1.1		<b>-2.7</b>		79	28	12	13
194724	<i>SABC</i> (CysA), chl. SO <sub>4</sub> <sup>2-</sup> transport system	<b>2.4</b>		-0.3		<b>-2.8</b>		20	4	3	3
184419	<i>APK</i> , APS kinase (chl.)	0.4	1.3	0.5	-0.7	<b>-1.7</b>	<b>-1.6</b>	39	29	12	41
131444	<i>APR</i> (MET16), APS sulforeductase (chl.)	0.5	1.4	-0.9	0.8	<b>-2.3</b>	-0.6	285	197	57	105
206154	<i>SIR1</i> , ferredoxin-sulfite reductase (chl.)	<b>2.0</b>		-1.1		<b>-2.3</b>		124	31	26	14
205485	<i>SIR2</i> , ferredoxin-sulfite reductase	<b>2.2</b>		-0.7		<b>-2.2</b>		597	126	132	80
189320	<i>OASTL4a</i> , O-acetylserine(thiol)-lyase (chl.)	<b>2.9</b>	<b>5.6</b>	-0.3	-0.3	<b>-5.0</b>	<b>-6.6</b>	976	128	31	103
205985	<i>SAT1a</i> , serineacetyl transferase (chl.)	<b>5.1</b>	<b>4.2</b>	-0.7	0.1	<b>-4.3</b>	<b>-4.3</b>	794	23	40	14
24268	<i>CGS1</i> (METB), cystathionine g-synthase (chl.)	<b>-1.6</b>		-0.6		1.4		23	71	63	49
127384	<i>THS1</i> , Thr synthase (chl.)	-1.4		-0.7		-0.3		54	138	42	84
196483	<i>METC</i> , cystathionine b-lyase (chl.)	-0.5		-0.5		1.3		24	34	57	24
154307	<i>METE</i> , Met synthase	-		-0.6		0.2		46	606	54	395
76715	<i>METH1</i> , Met synthase (cobalamin)	-		-0.9		-0.2		33	134	29	72
195332	<i>METH2</i> , Met synthase (cobalamin)	-		-0.8		1.1		30	211	65	123
182408	<i>METM</i> , S-adenosylmethionine synthetase	-		-0.2		-0.3		526	2288	420	1976
129593	<i>SAH1</i> , S-adenosylhomocysteine hydrolase	-		-0.4		-0.9		152	1298	83	952
183928	<i>DCA1</i> , S-adenosylmethionine decarboxylase	-		-0.3		1.5		54	184	149	146
206104	<i>AOT2</i> , a.a/polyamine transporter (p)	<b>4.1</b>	<b>3.6</b>	<b>-5.5</b>	<b>-2.4</b>	<b>-9.8</b>	<b>-6.3</b>	36	2	0	0
206105	<i>AOT4</i> , a.a/polyamine transporter (p)	<b>3.9</b>	<b>4.0</b>	<b>-8.0</b>	<b>-8.8</b>	<b>-13.0</b>	<b>-13.1</b>	311	21	0	0
185190	<i>THI4a</i> , thiazole biosynthetic enzyme	-		-0.1		<b>-2.8</b>		203	1859	29	1682
196899	<i>THI4b</i> , THI4 regulatory protein	-		-0.1		<b>-2.8</b>		188	1715	27	1557
192720	<i>THICa</i> , hydroxymethylpyrimidine phosphate synthase	-		-0.4		<b>-3.3</b>		6	95	1	74
196900	<i>THICb</i> , hydroxymethylpyrimidine phosphate synthase	-		-0.4		<b>-3.4</b>		6	92	1	72
142289	<i>BIOB1</i> , biotin synthase	-		1.1		-0.5		2	10	1	23
97943	<i>BIOB2</i> , biotin synthase	-		-0.1		<b>-1.6</b>		23	86	8	77

(Continued)

Table 3. (continued).

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type -S/ Wild Type +S		<i>snrk2.1</i> /Wild Type (+S)		<i>snrk2.1</i> /Wild Type (-S)		Wild Type	Wild Type +S	<i>snrk2.1</i> -S	<i>snrk2.1</i> +S
		S	Q	S	Q	S	Q	-S	+S	-S	+S
181975	<i>GSH1</i> , γ-glutamylcysteine synthetase (chl.)	0.1		0.0		0.1		35	33	38	34
189020	<i>GSH2</i> , glutathione synthetase (chl.)	-0.3		-0.3		-1.4		20	25	8	20
170508	<i>GTP1</i> , γ-glutamyl transpeptidase	0.2		0.7		0.7		25	23	41	37
Redistribution and recycling of S											
143892	<i>CDO1</i> , Cys dioxygenase	<b>4.7</b>	<b>4.5</b>	-1.4	<b>-4.0</b>	<b>-7.3</b>	<b>-8.4</b>	138	5	1	2
127464	<i>TAUD1</i> , taurin dioxygenase	<b>7.1</b>	<b>5.6</b>	<b>-3.5</b>	<b>-5.3</b>	<b>-11.7</b>	<b>-12.5</b>	1417	10	0	1
77600	<i>TAUD2</i> , taurin dioxygenase	<b>5.2</b>	<b>6.6</b>	-0.4	0.9	<b>-5.1</b>	<b>-6.4</b>	632	17	19	13
59800	<i>SUOX1</i> , sulfite oxidase, mit.	<b>2.0</b>		-0.2		<b>-1.8</b>		149	36	42	31
183511	<i>RDP3</i> , small rothhanase	<b>4.1</b>	<b>4.4</b>	-0.8	<b>-2.7</b>	<b>-4.0</b>	<b>-6.6</b>	369	21	24	12
132010	<i>TST</i> , mercaptopyruvate sulfurtransferase	0.2		0.4		-1.5		22	19	8	26
167884	<i>CSD2 (NIFS2)</i> , Cys desulfurase	<b>1.6</b>		-0.2		<b>-2.0</b>		43	14	10	12
98369	D-Cys desulfhydrase	<b>2.2</b>		0.2		<b>-2.5</b>		37	8	7	9
Lipid metabolism											
144554	Esterase/lipase/thioesterase	<b>3.1</b>		-0.5		<b>-1.6</b>		65	7	21	5
27658	<i>SQD1</i> , UDP-sulfoquinovose synthase	<b>1.7</b>		-0.9		<b>-4.1</b>		336	102	19	57
206163	<i>SQD2</i> , sulfolipid synthase	<b>1.5</b>		-0.3		<b>-5.1</b>		53	19	2	16
196477	<i>LPB1</i> , low phosphate bleaching	<b>4.1</b>	<b>4.2</b>	<b>-1.6</b>	-0.7	<b>-5.5</b>	<b>-5.8</b>	1711	98	39	33
Photosynthesis											
Stress-related genes											
148916	<i>ELIP3</i> , chlorophyll <i>a/b</i> binding protein	<b>2.9</b>		0.0		0.2		45	6	50	6
184724	<i>LHCSR1</i> , light-harvesting stress-related protein (chl.)	<b>1.8</b>	<b>2.2</b>	0.6	1.2	-1.3	<b>-2.1</b>	631	181	255	277
184731	<i>LHCSR2</i> , light-harvesting stress-related protein (chl.)	<b>2.2</b>	<b>4.6</b>	1.4	<b>5.9</b>	0.5	<b>2.0</b>	262	58	366	158
184730	<i>LHCSR3</i> , light-harvesting stress-related protein (chl.)	<b>2.4</b>	<b>5.5</b>	1.2	<b>7.0</b>	0.1	<b>3.0</b>	680	133	743	312
Major light-harvesting proteins of PSII											
185533	<i>LHCBM1</i> (chl.)	<b>2.9</b>		0.0		0.0		42	6	42	6
184067	<i>LHCBM2</i> (chl.)	<b>1.8</b>	<b>2.2</b>	0.6	1.2	<b>-1.6</b>	<b>-2.1</b>	596	174	202	257
195162	<i>LHCBM3</i> (chl.)	<b>2.1</b>	<b>4.6</b>	1.4	<b>5.9</b>	0.4	<b>2.0</b>	244	56	315	148
191690	<i>LHCBM4</i> (chl.)	<b>2.3</b>	<b>5.5</b>	1.1	<b>7.0</b>	-0.1	<b>3.0</b>	611	128	564	275
184775	<i>LHCBM5</i> (chl.)	-0.6		0.2		-1.1		5425	8019	2584	9327
184490	<i>LHCBM6</i> (chl.)	-0.2		0.2		-0.6		2705	3188	1740	3774
184071	<i>LHCBM7</i> (chl.)	<b>-1.7</b>		0.5		-0.2		676	2262	606	3112
205752	<i>LHCBM8</i> (chl.)	-0.3		0.3		<b>-2.2</b>		747	951	163	1160
184479	<i>LHCBM9</i> (chl.)	-1.4		0.6		<b>-2.6</b>		564	1485	95	2292
Carbon metabolism											
196354	<i>SHMT2</i> , Ser hydroxymethyltransferase 2	<b>-3.0</b>		-0.4		0.5		31	254	43	196
101528	6-Phosphogluconolactonase-like protein	<b>2.0</b>		-0.2		-1.3		401	100	160	85
163301	<i>GLD2-C</i> , glucose-6-phosphate dehydrogenase	<b>2.3</b>		-0.3		<b>-2.1</b>		77	16	18	13
192597	<i>GND1a</i> , 6-phosphogluconate dehydrogenase (decarboxylating) (chl.)	<b>1.9</b>		-0.2		<b>-1.6</b>		308	83	101	74
158911	<i>GND1b</i> , 6-phosphogluconate dehydrogenase (decarboxylating) (cyt.)	<b>1.9</b>		-0.2		<b>-1.6</b>		313	86	102	74
163238	<i>RIBA</i> , bifunctional GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase	<b>1.7</b>		-0.7		<b>-2.4</b>		83	26	15	16
196834	<i>CAH9</i> , carbonic anhydrase (cyt?)	<b>5.0</b>		-1.1		<b>-6.8</b>		1952	62	18	28

(Continued)

**Table 3.** (continued).

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type –S/ Wild Type +S		<i>snrk2.1</i> /Wild Type (+S)		<i>snrk2.1</i> /Wild Type (–S)		Wild Type	Wild Type +S	<i>snrk2.1</i> –S	<i>snrk2.1</i> +S
		S	Q	S	Q	S	Q	–S	+S	–S	+S
Folate metabolism											
206121	<i>MTDH/MTCH2</i> , methylenetetrahydrofolate dehydrogenase/methylenetetrahydrofolate cyclohydrolase	<b>–1.6</b>		–0.2		<b>–1.5</b>		13	40	5	35
127560	<i>FTFHS</i> , 10-formyltetrahydrofolate synthetase	<b>–2.5</b>		0.3		<b>–1.8</b>		10	59	3	71
30522	<i>ADCL1</i> , aminotransferase related to 4-amino-4-deoxychorismate lyase	<b>–1.7</b>		–0.2		0.2		23	73	26	65
111330	<i>MTHFR</i> , 5,10-methylenetetrahydrofolate reductase	<b>–3.1</b>		–0.1		0.1		11	92	12	86
182461	<i>GGH1</i> , $\gamma$ -glutamyl hydrolase	<b>3.9</b>		–0.6		<b>–6.2</b>		36	2	0	2
Extracellular/cell wall proteins											
168785	<i>HAP1</i> , vanadium haloperoxidase	<b>7.9</b>	<b>7.8</b>	–0.2	<b>–2.0</b>	<b>–8.4</b>	<b>–11.7</b>	739	3	2	3
143696	<i>HAP2</i> , vanadium haloperoxidase	<b>8.4</b>	<b>7.1</b>	–0.1	–0.6	<b>–11.6</b>	<b>–11.5</b>	1044	3	0	3
182794	<i>HAP3</i> , vanadium haloperoxidase	<b>7.1</b>	<b>6.5</b>	<b>–3.2</b>	–1.0	<b>–11.1</b>	<b>–7.2</b>	1456	10	1	1
130684	<i>ECP76</i> , extracellular protein	<b>10.7</b>	<b>10.3</b>	–0.8	–0.4	<b>–13.1</b>	<b>–11.8</b>	899	1	0	0
137329	<i>ECP88</i> , extracellular protein	<b>13.4</b>		<b>–2.3</b>		<b>–15.9</b>		2295	0	0	0
119420	<i>ECP61</i> , extracellular protein	<b>8.5</b>		–0.6		<b>–11.0</b>		268	1	0	0
194201	<i>ECP56</i> , extracellular protein	<b>8.6</b>		<b>–4.2</b>		<b>–13.0</b>		317	1	0	0

Fold change (log<sub>2</sub> values) in transcript levels under specified conditions as determined by RNA-seq (S) and, when available, by qRT-PCR (Q). Values above 1.5 or below –1.5 are marked in bold type. Some genes for which expression is not affected by –S conditions are included for informative purposes. Putative subcellular localizations are denoted in parentheses after the transcript names as cyt, cytosolic; mit, mitochondria; chl, chloroplast; and ext, extracellular. ID, JGI protein accession number; p, putative; nd, not enough high-quality data for statistical tests.

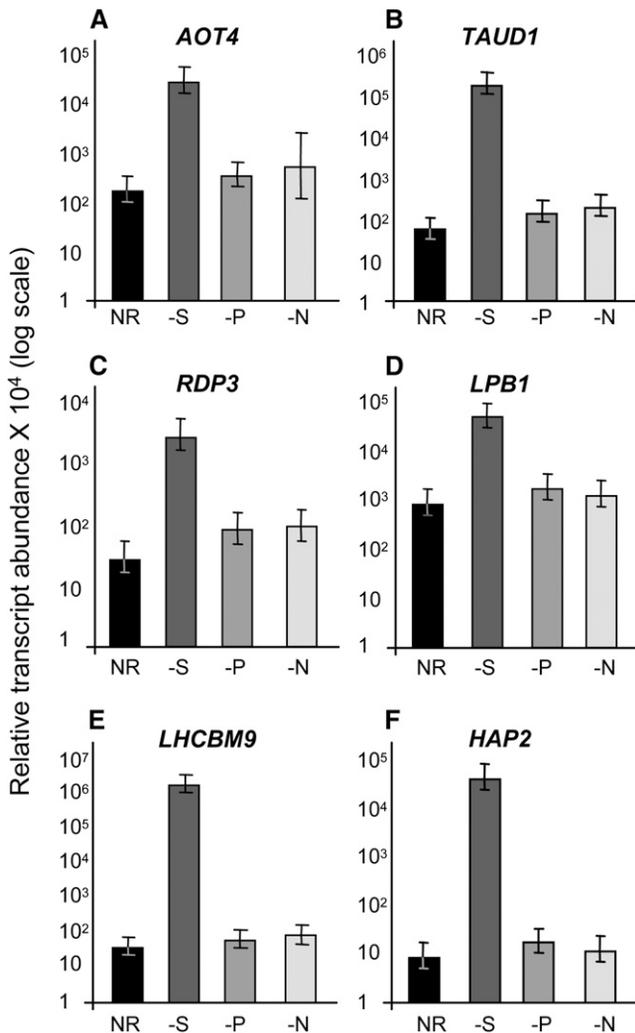
<sup>a</sup>Transcript abundance obtained from RNA-seq data is indicated as RPKMs (see Methods).

These results demonstrate that –S conditions elicit a massive rise in *LHCBM9* mRNA that is specifically controlled by previously identified S regulatory elements and that this rise is concurrent with a loss in transcripts encoding *LHCBM1-8*. The possible reduction in *LHCBM1-8* polypeptide synthesis may be balanced by an increase in *LHCBM9* synthesis. These results could also explain the photosynthetic phenotypes of the *snrk2.1* and *sac1* mutants; a severe reduction in *LHCBM1-8* transcripts without the concomitant increase in *LHCBM9* in S-deprived *snrk2.1* may compromise the overall integrity of PSII and elicit elevated production of reactive oxygen species (ROS), and more specifically the photoproduction of singlet oxygen (<sup>1</sup>O<sub>2</sub>). This hypothesis is supported by the findings that transcripts encoding proteins important for ameliorating the consequences of <sup>1</sup>O<sub>2</sub> production are elevated in the S-starved *snrk2.1* mutant, but not in S-starved wild-type cells (see section: ROS Detoxification and Cellular Redox). Elevated <sup>1</sup>O<sub>2</sub> levels can cause photooxidative damage, reduce cell viability, and elicit programmed cell death; genes encoding proteins potentially involved in apoptosis are elevated in *snrk2.1* but not in wild-type cells under –S conditions (see section: Apoptosis and Autophagy). A model depicting correlative relationships between photosynthetic electron flow, *LHCBM9*, <sup>1</sup>O<sub>2</sub>, and programmed cell death during S deprivation in the mutant and wild-type cells is shown in Figure 9; more work is needed to substantiate this model.

Similar to *LHCBM9* mRNA, transcripts encoding four unique LHC polypeptides (*LHCSR1*, *LHCSR2*, *LHCSR3*, and *ELIP3*) that are not associated with light-harvesting function markedly increased in cells experiencing –S conditions. *LHCSR2* and *LHCSR3* transcript levels increased to a similar extent in the S-starved wild type and the *snrk2.1* mutant, although the levels of these transcripts during nutrient-replete growth were also somewhat higher in the mutant cells. By contrast, like *LHCBM9* mRNA, the increase in the *LHCSR1* transcript observed in wild-type cells under –S conditions was not observed in the *snrk2.1* mutant (Table 3). *Chlamydomonas* LHCSR polypeptides have no homologs in vascular plants and have been associated with photoprotection during stress conditions (Richard et al., 2000; Elrad and Grossman, 2004; Peers et al., 2009), and, in a previous study, the level of *LHCSR2* mRNA was shown to increase in cells deprived of S (Zhang et al., 2004).

### Carbon Metabolism

Transcripts encoding key proteins of the Calvin-Benson cycle (reductive pentose phosphate pathway), the glyoxylate cycle, and the C<sub>4</sub> dicarboxylic acid cycle declined in both wild-type and *snrk2.1* mutant cells experiencing S deprivation; this decline was generally more severe in mutant than wild-type cells (see



**Figure 6.** Abundances of Six Selected Transcripts under Different Nutrient Deprivation Conditions.

RNA samples were extracted from wild-type cells for qRT-PCR analysis at 6 h following their transfer to the different media (NR, nutrient-replete; -S, S-depleted; -P, phosphorus-depleted; -N, nitrogen-depleted). Levels of individual transcripts noted in (A) to (F) are given as relative fold abundance with respect to the housekeeping control gene (*CBLP*) and then multiplying relative target gene abundances by a factor of 10<sup>4</sup>. Each of the values was from at least three replicates, with the error bars representing 1 SD. Graphics are in log scale; note the different scales for the different analyses.

Supplemental Data Set 2B online). Moreover, transcripts for most proteins required for folate biosynthesis also declined, while the transcript encoding  $\gamma$ -glutamyl hydrolase (GGH1), an enzyme involved in folate catabolism (Scott et al., 2000), increased following the imposition of -S conditions (Table 3). Various folate derivatives, including tetrahydrofolates (THFs), are one-carbon donors and acceptors required for many cellular processes, including the synthesis of Gly from Ser, methylation of dUMP to form dTMP, the biosynthesis of pantothenate, purine biosynthesis, formylation of fMet-tRNA, and Met biosynthesis. In

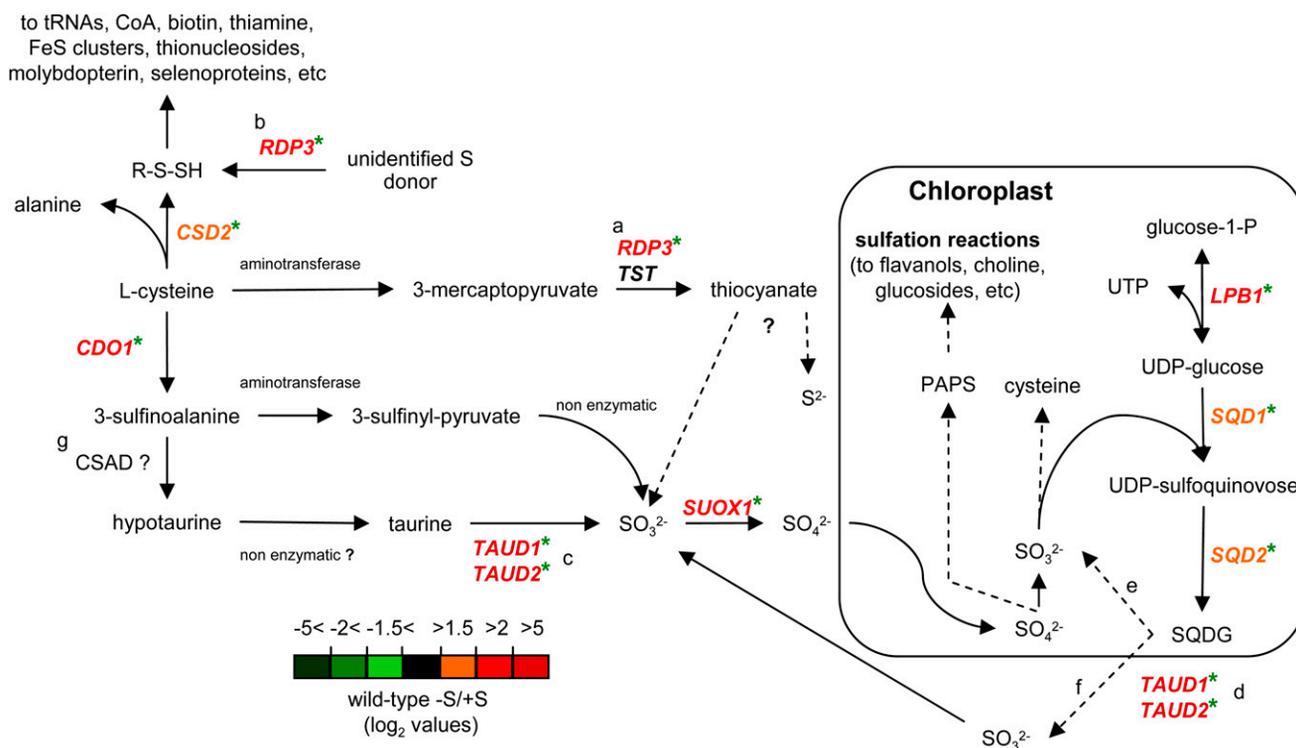
plants, the majority of folate (5-methylTHF) is used for the synthesis of Met/SAM (Hanson and Roje, 2001). The decline in transcript levels for enzymes involved in THF synthesis parallels the decline in Met levels (Bolling and Fiehn, 2005) and in the levels of those transcripts encoding Met/SAM cycle enzymes (Table 3) in S-deprived *Chlamydomonas*. Furthermore, since THF is required for the Ser-Gly interconversion reaction of the photorespiration/glyoxylate cycle, it is likely that photorespiratory activity declines in S-deprived cells. A reduction in photorespiration is also suggested by a decrease in the level of the transcript encoding Ser hydroxymethyltransferase (SHMT2) (Table 3), a key enzyme of the pathway that requires a THF cofactor. Similarly, a decline in photorespiration and folate metabolism as consequence of reduced SAM levels in S-depleted *Arabidopsis* plants has been proposed (Nikiforova et al., 2005b), although no empirical data were presented to substantiate this conjecture.

Transcripts encoding proteins of the oxidative pentose phosphate (OPP) pathway increased in S-deprived *Chlamydomonas* cells, which may be reflected by accumulation of OPP metabolites (Bolling and Fiehn, 2005). In plants, an increased flux of metabolites through OPP is stimulated by stress conditions and is linked to ROS detoxification (Couee et al., 2006). Alternatively, elevated OPP activity may provide cells with an additional capacity to generate NADPH under conditions of diminished photosynthetic activity. The increase in transcripts encoding OPP enzymes was less in *snrk2.1* relative to wild-type cells (Table 3). Furthermore, S deprivation elicited a rise in the level of a transcript encoding the bifunctional GTP cyclohydrolase II/3,4-dihydroxy-2-butanone-4-phosphate synthase (RIBA). This enzyme uses ribulose-5-P and GTP to catalyze the first step of riboflavin biosynthesis. Elevated RIBA levels may route excess ribulose-5-P formed as the final product of OPP into riboflavin, FMN, and FAD.

A number of other interesting observations have been noted that suggest significant modifications in carbon metabolism upon exposure of *Chlamydomonas* to -S conditions. S-starved algal cells accumulate high levels of starch (Zhang and Melis, 2002), suggesting a strong redirection of fixed carbon into storage, most likely because the cells cease dividing and have reduced anabolic metabolism. However, there is no significant increase in transcripts for enzymes involved in starch synthesis. Similar results have been observed for rice (*Oryza sativa*) plants (Lunde et al., 2008). We also noted that transcripts encoding the putative carbonic anhydrases CAH7 and CAH9 (Moroney and Ynalvez, 2007) increased in S-deprived cells; the CAH9 transcript exhibited a 31-fold increase that was SNRK2.1 dependent (Table 3). A CAH gene was also found to be highly upregulated when *Arabidopsis* plants were deprived of S (Maruyama-Nakashita et al., 2003). Although intriguing, the role of these CAHs during acclimation of organisms to -S conditions is unexplored and not easily explained at this time.

### Extracellular Proteins

Extracellular arylsulfatases (ARSs) and putative extracellular cell wall polypeptides (ECP76 and ECP88), as well as transcripts encoding these peptides, specifically accumulated when *Chlamydomonas* was deprived of S (de Hostos et al., 1988; Takahashi



**Figure 7.** Proposed Pathways for Redistribution and Recycling of S.

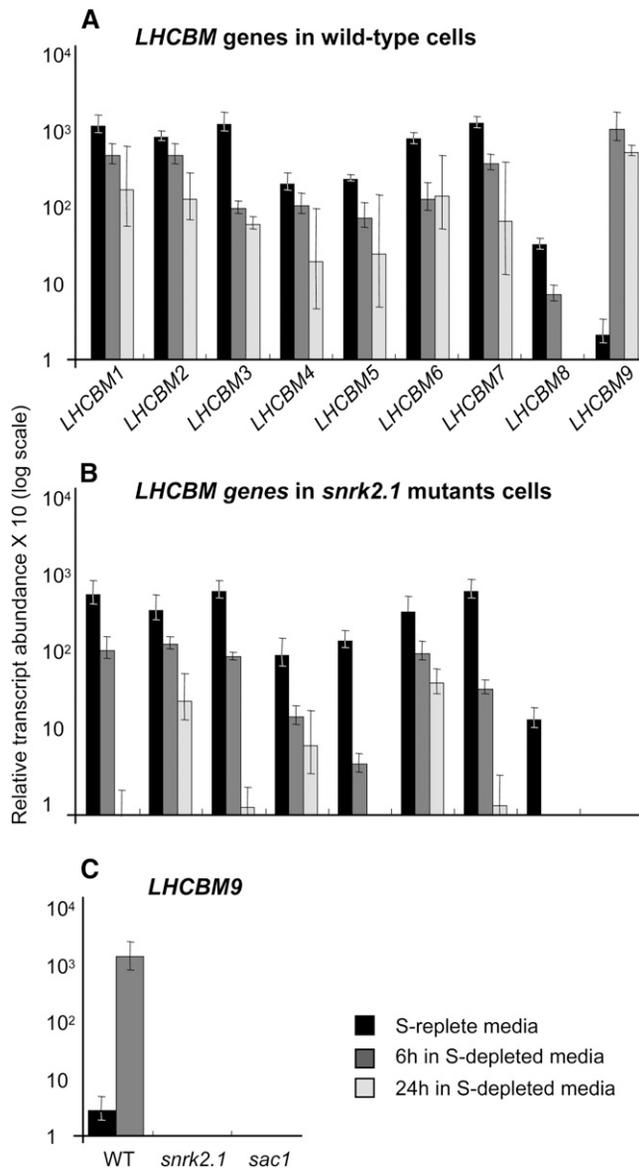
The “a” and “b” denote two possible functions for RDP3; “c” and “d” denote two possible functions for TAUD1 and TAUD2; “e” and “f” denote two possible subcellular compartments in which SQDG degradation may occur; “g”, no gene encoding sulfinoalanine decarboxylase (CSAD) was identified on the *Chlamydomonas* genome. Other details are as in Figure 5. PAPS, adenosine 3'-phosphate 5'-phosphosulfate;  $S^{2-}$ , sulfide.

et al., 2001). In this study, we identified five new transcripts encoding putative extracellular/cell wall proteins that increase during S deprivation; two encode ECP-like proteins, while the remaining three encode putative extracellular vanadium haloperoxidases (HAPs) (Table 3).

The two transcripts encoding novel ECP-like proteins, ECP61 and ECP56 (protein IDs 119420 and 194201, respectively; the deduced molecular masses follow the ECP designation), rose by up to 268- and 317-fold, respectively, in cells experiencing  $-S$  conditions (Table 3). The increase in all ECP transcripts was completely abolished in the *snrk2.1* mutant. Like ECP88 and ECP76, these new putative proteins have features similar to those of cell wall, Hyp-rich glycoproteins with very low S-containing amino acid content (Table 4). While the synthesis of the S-responsive ECPs is likely a means to economize on the use of limited available S, it may also have significant physiological consequences (e.g., rigidity of the wall and protection of the cell body during conditions of growth limitation).

*HAP1*, *HAP2*, and *HAP3* transcripts increased by up to 246-, 348-, and 145-fold, respectively, in wild-type cells experiencing S deprivation; these increases were strictly dependent on SNRK2.1 (Table 3). Interestingly, based on microarray data, the *HAP2* mRNA increased within 1 h of transferring cells to  $-S$  conditions (see Supplemental Data Set 2A online), but did not increase in response to  $-P$  or  $-N$  conditions (Figure 6F). HAPs, often closely related to acid phosphatases and mostly localized

outside of the cell body, can use  $H_2O_2$  to catalyze the oxidative halogenation of organic compounds. These enzymes may also be responsible for extracellular peroxidase activity previously noted in *Chlamydomonas* (Torkelson et al., 1995). However, HAPs generally have broad substrate specificity and in addition to halide oxidation, may catalyze sulfoxidation, epoxidation, and phosphodiesterase reactions (Hemrika et al., 1997; Dembitsky, 2003). Because of their broad potential reaction scope and substrate specificity, it is difficult to assign HAPs a specific physiological function. Putative roles for these enzymes under  $-S$  conditions would include (1) participation in the biosynthesis/degradation of cell wall polyphenols (Vreeland et al., 1998; Ortiz-Bermudez et al., 2007) in reactions linked to  $H_2O_2$  production (Slesak et al., 2007); such activities may alter cell wall architecture. (2) They function as an extracellular phosphatase in S-deprived cells. Such an activity, which is both SNRK2.1 and SAC1 dependent, has been previously observed (Moseley et al., 2009), although its physiological significance is not known. (3) They contribute to the conversion of sulfides to sulfoxides, possibly the initial step in a novel mechanism for scavenging external sulfides. Sulfides frequently present in the environment include Met, SAM, cystathionine, dimethylsulfide, and 3-methylpropionate. The latter two compounds are intermediary metabolites in dimethylsulfoniopropionate degradation, a potentially abundant, ecologically important molecule in the environment (Bentley and Chasteen, 2004).



**Figure 8.** Abundances of *LHCBM* Transcripts Based on qRT-PCR.

(A) and (B) RNA samples were extracted from the wild type (A) and the *snrk2.1* mutant (B) and analyzed for all nine of the *LHCBM* transcripts. (C) Comparisons of *LHCBM9* expression levels in wild-type, *snrk2.1*, and *sac1* cells. Levels of individual transcripts are given as relative fold abundances with respect to the housekeeping gene (*CBLP*) and then multiplying the relative target gene abundances by a factor of 10. Graphics are in log scale. Each of the values was from at least three replicates, with the error bars representing 1 SD. None of the values obtained were much below 1, and those values that were below 1 are represented as 0 on the graph.

### S Sparing at the Proteome Level

As proposed previously (Mazel and Marliere, 1989; Takahashi et al., 2001; Pereira et al., 2008), replacing high S content proteins with others that have low S content represents an S-sparing response that would help conserve cellular S re-

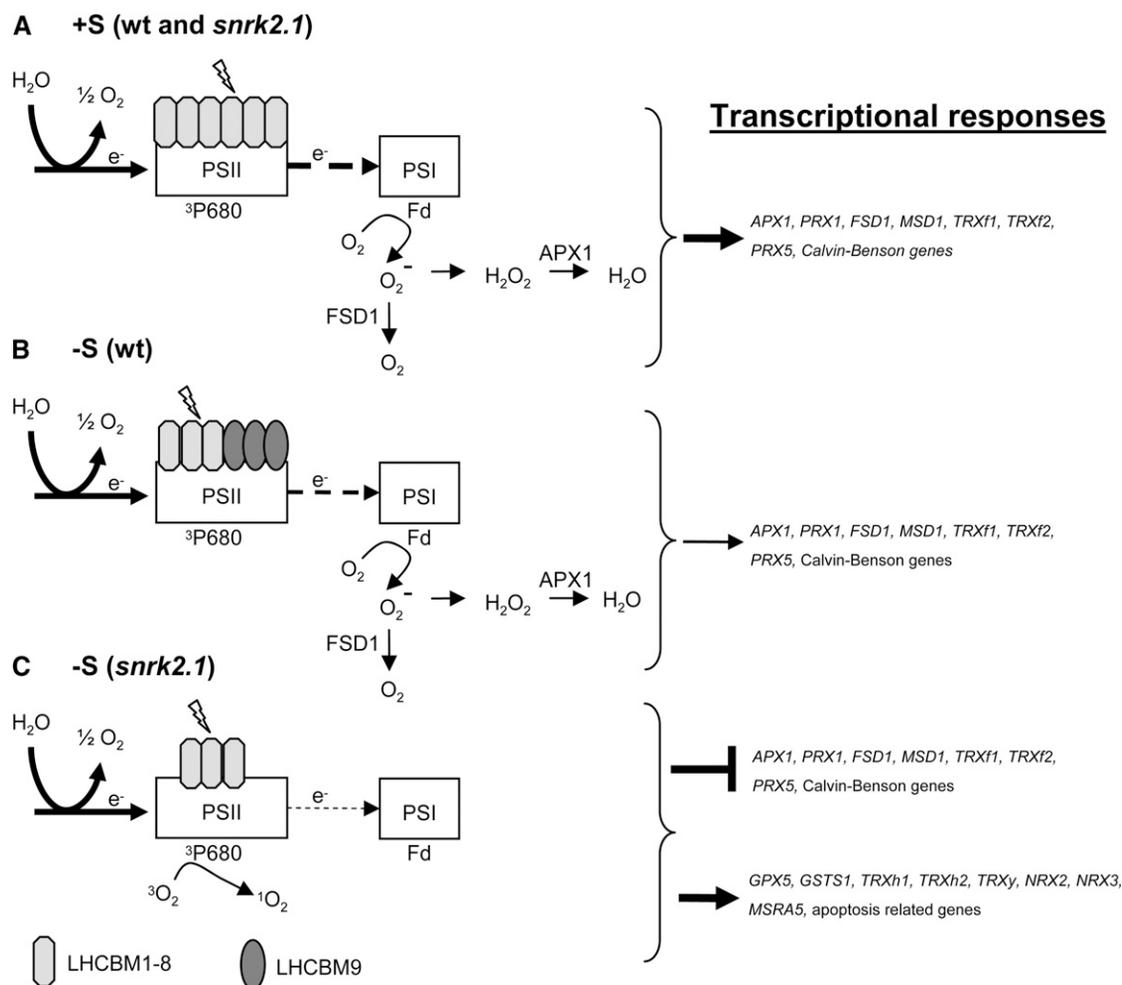
sources. In *Chlamydomonas*, the swapping of proteins with high S content for ones with low S content was previously proposed with respect to the putative cell wall polypeptides ECP76 and ECP88 (Takahashi et al., 2001) and, as discussed above, can now be extended to ECP61 and ECP56 (Table 4). Interestingly, the *LHCBM9* polypeptide also has a low content of S-containing amino acids. There are three invariant S-containing amino acids present in all mature *LHCBM* polypeptides except *LHCBM9* (Figure 10). Hence, mature *LHCBM9* has at least three fewer S-containing amino acids than the other *LHCBM*s (Table 4). Since *LHCBM* polypeptides are abundant cellular constituents, substitution of S-rich *LHCBM1-8* with the S-poor *LHCBM9* would allow for significant S conservation. An analogous situation has been described for cyanobacterial LHC proteins (Mazel and Marliere, 1989). Moreover, many other *Chlamydomonas* transcripts that accumulated during -S conditions encode proteins that contain relatively few S-containing amino acids (e.g., ARSSs, HAPs, TAUDs, among others), and the average S amino acid content of proteins encoded by the transcripts that accumulate during S deprivation is lower (3.32%) than that of the average *Chlamydomonas* proteome (4.19%) (Table 5). This decrease in S amino acid content is even more dramatic if the percentages of these proteins are weighted by the abundance of their corresponding transcripts in S-deprived cells (2.33%) (Table 5). The SNRK2.1 kinase is critical for this S-sparing response (Table 5). A similar S-sparing phenomenon has been described for yeast (Fauchon et al., 2002). While these data are not quantitative with respect to the relative amount of S-containing amino acids used for protein biosynthesis (absolute mRNA levels and/or identification and quantification of the proteins in the proteome would be required), they do suggest that proteins that accumulate under -S conditions tend to have a reduced S amino acid content, which likely helps optimize the use of the limiting resource.

### Other Significant Processes

S deprivation also elicited an increase in transcripts encoding several putative plasma membrane transporters and 83 RNAs and/or proteins of unknown function (see Supplemental Data Set 2A online), along with a decrease in the levels of transcripts for proteins associated with amino acid biosynthesis, lipid metabolism, some transport processes and translocons, tRNA biosynthesis, chaperones, translation (especially chloroplast ribosomal proteins), and transcription (see Supplemental Data Set 2B online). A decrease in the levels of mRNAs encoding proteins associated with a variety of metabolic processes may reflect an overall slowing of metabolism as the cells stop dividing.

### S Deprivation Responses Specific to the *snrk2.1* Lesion

Transcripts of wild-type and *snrk2.1* mutant cells that were responsive to S deprivation were categorized based on their transcriptome profiles (Figure 3; see Supplemental Data Set 2 online). Those genes that appear to be under SNRK2.1 control (directly or indirectly) are in categories A1 and B1 in Figure 3 and Supplemental Data Set 2 online. Transcripts that change in abundance independently of SNRK2.1 are in categories A2-A3



**Figure 9.** Proposed Model for How the *snrk2.1* Lesion Affects Photosynthesis and ROS Production Processes.

Wild-type (wt) and mutant cells in nutrient-replete medium **(A)** perform photosynthesis at optimal levels with the consequent production of superoxide ( $O_2^-$ ) and  $H_2O_2$ . Under these conditions, transcripts encoding proteins related to  $O_2^-$  and  $H_2O_2$  detoxification are relatively abundant, along with transcripts encoding Calvin-Benson-Bassham cycle-related enzymes. In wild-type cells that experience S deprivation **(B)**, photosynthetic activity declines (most photosynthesis, chlorophyll biosynthesis, and Calvin-Benson-Bassham cycle-related transcripts decline), and the LHCBM1-8 proteins are partially replaced by LHCBM9. The production of  $O_2^-$  and  $H_2O_2$  may be diminished because of reduced PSII activity and electron flow to PSI. Transcripts encoding proteins related to the detoxification of these specific ROS also decline. In the *snrk2.1* mutant experiencing S deprivation **(C)**, LHCBM9 is not synthesized and the PSII antenna may become reduced in size. This may impair the proper assembly of PSII and its capacity for photoprotection, triggering the production of singlet oxygen ( $^1O_2$ ) and of genes specifically associated with  $^1O_2$  accumulation. Aberrant PSII architecture,  $^1O_2$  accumulation, and severe S deficiency could dramatically impair PSII function in the *snrk2.1* mutant, resulting in a marked decline in electron flow to PSI. This decline in the photosynthetic activity would result in a decrease in the levels of transcripts encoding Calvin-Benson-Bassham cycle enzymes as well as those related to  $O_2^-$  and  $H_2O_2$  accumulation. Accumulation of  $^1O_2$  may also trigger expression of genes potentially involved in the initiation of apoptosis (listed in Table 6). Thickness of the arrow between PSII and PSI depicts a qualitative representation of the electron flow rate. See “ROS Detoxification and Cellular Redox” section for more details.

and B2-B3. However, there is also a large set of transcripts that changed in abundance in S-deprived *snrk2.1* but not in S-deprived wild-type cells. Included in these mutant-specific changes, presented as categories C1 and C2, are transcripts encoding proteins associated with oxidative stress, proteolysis, and apoptosis. Some genes encoding proteins critical for the biosynthesis of the S-containing cofactor CoA also increased in S-deprived mutant relative to S-deprived wild-type cells (Table 6). All these changes potentially reflect the severe S deficiency

and the highly aberrant physiological state exhibited by the mutant because of its inability to acclimate to S-limiting conditions.

#### ROS Detoxification and Cellular Redox

Transcripts encoding GSH-peroxidase 5 (GPX5), GSH-transferase 1 (GSTS1), and thioredoxin h1 (TRXh1) proteins increased in the *snrk2.1* mutant relative to wild-type cells following the

**Table 4.** S-Containing Amino Acid Content of Cell Wall and PSII Light-Harvesting Proteins

ID	Protein	Length	C+M <sup>a</sup>	C+M (%) <sup>b</sup>
Cell wall proteins				
130684	ECP76, cell wall protein	583	1 (+1)	0.2
137329	ECP88, cell wall protein	595	0 (+4)	0.0
119420	ECP61, cell wall protein	577	1 (+1)	0.2
194201	ECP56, cell wall protein	532	1 (+1)	0.2
195768	GP2, cell wall protein	1182	48 (+1)	4.1
196402	PHC2, cell wall protein	505	36 (+2)	7.1
196403	PHC3, cell wall protein	443	25 (+1)	5.6
195824	VSP4, cell wall protein	993	43 (+2)	4.3
196399	PHC1, cell wall protein	479	36 (+2)	7.5
130359	VSP6, cell wall protein	951	46 (+1)	4.8
148333	PHC15, cell wall protein	184	10 (+3)	5.4
194264	PHC12, cell wall protein	407	28 (+3)	6.9
187643	PHC14, cell wall protein	566	42 (+1)	7.4
193780	GAS31, cell wall protein	507	37 (+1)	7.3
Peripheral PSII antennae proteins				
184479	LHCBM9	254	2 (+2)	0.8
185533	LHCBM1	256	5 (+1)	2.0
184067	LHCBM2	249	5 (+2)	2.0
195162	LHCBM3	257	5 (+1)	1.9
191690	LHCBM4	254	6 (+1)	2.4
184775	LHCBM5	268	5 (+2)	1.9
184490	LHCBM6	253	5 (+2)	2.0
184071	LHCBM7	249	5 (+2)	2.0
205752	LHCBM8	254	5 (+2)	2.0

The S-containing amino acid levels of the ECP and LHCBM9 proteins are compared with the LHCBM1-8 proteins and the 10 most abundant putative cell wall proteins (according to RNA-seq data).

<sup>a</sup>Total number of Cys and Met residues (C+M) for the mature proteins, with the numbers between parentheses indicating the C+M content in the putative signal peptides.

<sup>b</sup>C+M percentages are for the mature proteins.

imposition of S deprivation (Table 6); these proteins are known to mitigate the effects of ROS-elicited lipid peroxidation and DNA damage (Goyer et al., 2002; Sarkar et al., 2005; Dayer et al., 2008) by amelioration the toxic effects of <sup>1</sup>O<sub>2</sub> (Ledford et al., 2007). Furthermore, *TRXh1* and *GPX5* transcript accumulation in *Chlamydomonas* is either not affected or affected to a minor extent by elevated superoxide (O<sub>2</sub><sup>-</sup>) or H<sub>2</sub>O<sub>2</sub> concentrations (Lemaire et al., 1999; Sarkar et al., 2005; Ledford et al., 2007). Ledford et al. (2007) identified 14 transcripts that changed in abundance when *Chlamydomonas* cells experienced high <sup>1</sup>O<sub>2</sub> levels. These transcripts displayed a similar change in abundance in S-deprived *snrk2.1* but not in S-deprived wild-type cells (see Supplemental Table 1 online), suggesting that S deprivation of the mutant elicits sustained production of <sup>1</sup>O<sub>2</sub>. In photosynthetic organisms, <sup>1</sup>O<sub>2</sub> is mainly produced as a consequence of PSII electron flow, especially when the organisms are absorbing excess excitation energy, and is thought to be the main cause of photooxidative damage and cell death in plants (Apel and Hirt, 2004; Asada, 2006; Triantaphylides et al., 2008; Triantaphylides and Havaux, 2009). As proposed in the section "Photosynthesis" and in Figure 9, an aberrant light-harvesting architecture in *snrk2.1* during -S conditions (likely associated with a deficit in LHCBM9) may severely impact the activity of PSII and its

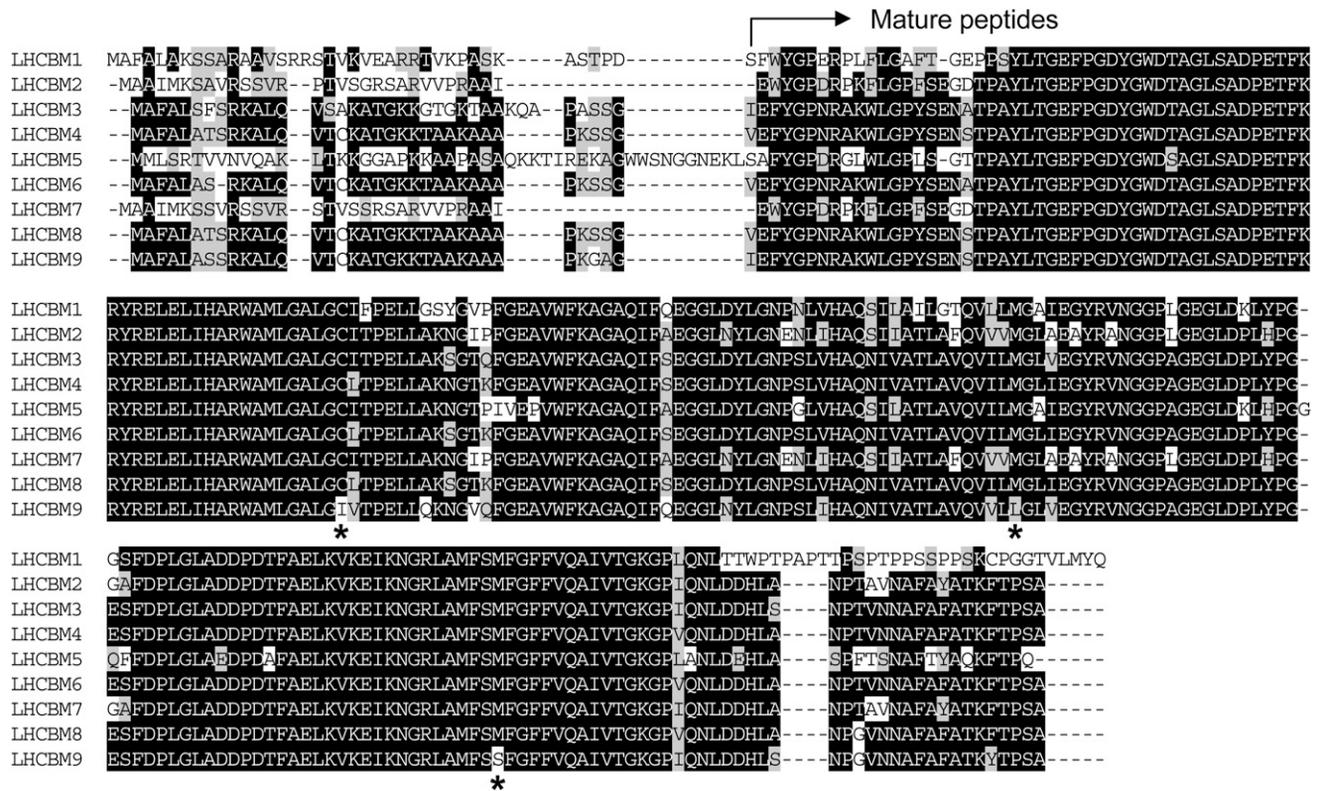
capacity for photoprotection, which in turn could cause elevated <sup>1</sup>O<sub>2</sub> production.

Interestingly, transcripts encoding proteins more specific to combating the lethal effects of superoxides and H<sub>2</sub>O<sub>2</sub>, including L-ascorbate peroxidases 1 and 3 (APX1 and APX3), peroxiredoxin 1 (PRX1), mitochondrial superoxide dismutase 1 (MSD1), and chloroplast superoxide dismutase 1 (FSD1), were slightly reduced in wild-type cells, but their reduction was dramatic in *snrk2.1* cells deprived of S (Table 6). These data suggest that the production of O<sub>2</sub><sup>-</sup> and H<sub>2</sub>O<sub>2</sub> declined during S deprivation, especially in the *snrk2.1* strain. The production of O<sub>2</sub><sup>-</sup> and H<sub>2</sub>O<sub>2</sub> are mostly generated as a consequence of peroxisome metabolism, respiration, and hyperreduction of PSI (Apel and Hirt, 2004; Asada, 2006). In wild-type cells during -S conditions, photosynthetic electron transport is downregulated (Wykoff et al., 1998), which would likely reduce the production of O<sub>2</sub><sup>-</sup> and H<sub>2</sub>O<sub>2</sub> because of reduced electron flow to PSI. In S-deprived *snrk2.1*, the flow of electrons to PSI would potentially be more severely diminished, as depicted in Figure 9.

Other proteins encoded by transcripts elevated exclusively in the *snrk2.1* mutant during S deprivation include TRXh2, TRXy, Met sulfoxide reductase 5 (MSRA5), glutaredoxin 2 (GRX2), and nucleoredoxins 2 and 3 (NRX2 and NRX3). By contrast, S deprivation of the mutant cells caused a decline in the levels of *PRX5*, *TRXf1*, and *TRXf2* transcripts; the level of *TRXf2* mRNA also declined in -S wild-type cells, although absolute levels were significantly lower in the mutant (Table 6). Elevated *MSRA5* transcripts levels may reflect protein damage by oxidative conditions. GRX activity is required for protein deglutathionylation; the role of this enzyme in glutathionylation is not understood (Rouhier et al., 2008). Many chloroplast proteins are nonenzymatically glutathionylated during oxidative stress (including TRXf, PRX1, and Calvin-Benson cycle enzymes), resulting in inhibition of their activities and/or protection from irreversible oxidation (Michelet et al., 2005, 2008). Thus, the high levels of the GRX2 may alter the glutathionylation state of many chloroplast proteins during -S conditions in the *snrk2.1* mutant. TRXf polypeptides, located in chloroplasts, are reductively activated by ferredoxin (Fd) and participate in light-dependent activation/deactivation of key carbon metabolism enzymes, especially those of the Calvin-Benson cycle. Decreased TRXf transcript levels during S deprivation may reflect a decrease in photosynthetic electron flow through PSI (Figure 9); this decrease was more pronounced in *snrk2.1* relative to the wild type (Table 6). Additionally, the redox state of TRXf is expected to decrease, which would alter its activity and the activities of its target proteins.

### Apoptosis and Autophagy

The S-deprived *snrk2.1* mutant exhibited an increase in the levels of six transcripts encoding proteins previously associated with apoptosis, as well as a decline in a transcript encoding the antiapoptosis protein (DAD1) (Table 6). There were also 33 transcripts for proteins related to autophagy and vesicular and vacuolar transport (including the exocyst complex, GTPases, R-SNARE and Q-SNARE family proteins, among others) that increased specifically in *snrk2.1* cells (and not in wild-type cells) in response to S deprivation (see Supplemental Data Set 2C1



**Figure 10.** Alignments of the Amino Acid Sequences of *Chlamydomonas* LHCBM Proteins.

The LHCBM9 polypeptide has substitutions at positions in which there are conserved S-containing amino acids in the other LHCBM proteins; these are Met157Leu, Met213Ser, and Cys101Ile; they are marked with an asterisk. The black and gray boxes indicate identical and similar amino acids, respectively. Alignments were performed using BioEdit 7.0.5.3 software.

online); plant cells experiencing apoptosis display intense autophagy, with the vacuole acting like an animal autophagosome (Williams and Dickman, 2008). Finally, numerous transcripts encoding proteins associated with proteolysis, nuclease activity, chromatin condensation, RNA binding, DNA binding, and lipid degradation also increased specifically in the *snrk2.1* mutant under -S conditions (see Supplemental Data Set 2C1 online). The proteins encoded by all of these transcripts are potentially involved in apoptosis/autophagy, which could explain the severe bleaching and rapid death of *snrk2.1* cells during -S conditions (González-Ballester et al., 2008). Apoptosis-like processes have been previously noted for *Chlamydomonas* cells exposed to UV-C irradiation (Moharikar et al., 2006, 2007) and for plants experiencing photooxidative damage and elevated levels of <sup>1</sup>O<sub>2</sub> (Triantaphylides et al., 2008; Williams and Dickman, 2008). Thus, these data support the hypothesis that aberrant light-harvesting architecture, elevated <sup>1</sup>O<sub>2</sub> production, and a loss of cell viability are all linked to the inability of the *snrk2.1* mutant to acclimate to -S conditions (Figure 9).

**Putrescine Biosynthesis**

The transcript for Orn decarboxylase (ODC2) increased by up to 10-fold specifically in the *snrk2.1* mutant during S deprivation (Table 6). ODC2 catalyzes the conversion of Orn to putrescine,

the precursor of the polyamines spermine and spermidine. Although Orn biosynthesis occurs mostly through the acetyl cycle of Arg, the increase in *ODC2* transcript levels in *snrk2.1* during -S conditions was accompanied by a decrease in most transcripts encoding enzymes involved in Arg biosynthesis; the only exception was that of the *ARG9* transcript, which encodes an *N*-acetylmethionine aminotransferase (NAOAT) (Table 6, Figure 11). An alternative route to Orn/polyamine synthesis involves Pro dehydrogenase (PRD), coupled with Orn aminotransferase activity (OAT) (Slocum, 2005) (Figure 11). In the *snrk2.1* mutant, *PRD* transcript levels increased under -S conditions, although the level of the *OAT1* transcript showed little change. Interestingly, several reports indicate that some NAOATs can also use the nonacetylated compounds Orn and Glu-5-semialdehyde as substrates (Billheimer et al., 1976; Heimberg et al., 1990; Ledwidge and Blanchard, 1999). Therefore, it is possible, although still speculative, that *Chlamydomonas* NAOAT (*ARG9*) also works as an OAT. Additionally, the level of the transcript encoding Glu-5-semialdehyde dehydrogenase (*GSD1*) increased in S-starved *snrk2.1* cells (Table 6, Figure 11). The coordinated activities of PRD, *GSD1*, *ARG9* (acting as an OAT), and *ODC2* may define an alternative pathway for the synthesis of putrescine from Pro and Glu in -S *snrk2.1* cells (Figure 11). This pathway could be predominant over the acetyl cycle of Arg under conditions in which levels of the S-containing CoA, a cofactor

**Table 5.** S-Amino Acid Content of *Chlamydomonas* Proteins

List of Transcripts <sup>a</sup>	C+M % (nwt) <sup>b</sup>	C+M% Weighed by Transcript Abundance <sup>b</sup>					
		+S (Wild Type)	–S (Wild Type)	+S ( <i>snrk2.1</i> )	–S ( <i>snrk2.1</i> )	+P (Wild Type)	–P (Wild Type)
All	4.19	4.55	4.32	4.36	4.54	4.20	4.25
–S up (wild type)	3.32	3.83	2.33	3.28	3.09	4.08	3.13
–S down (wild type)	4.05	3.93	3.89	3.79	4.24	3.64	3.49
–S up ( <i>snrk2.1</i> )	4.18	4.25	4.02	4.28	4.28	4.45	4.28
–S down ( <i>snrk2.1</i> )	3.98	3.66	3.68	3.57	3.64	3.64	3.62
–P up (wild type)	4.18	4.42	3.12 <sup>c</sup>	4.73	4.29	4.43	4.06
–P down (wild type)	4.18	4.61	4.48	4.37	4.07	4.32	4.30

Phosphorous deprivation (–P) was used as internal control. nwt, data not weighted by transcripts abundance.

<sup>a</sup>The list of transcripts that accumulate (up) or decline (down) under the different conditions were selected as described in Figure 3. Subsequently, the genes encoding proteins with no significant homology to a known protein (unknown genes) were filtered out because many of them represent inaccurate v.3.1 gene models and would introduce a strong bias into the data.

<sup>b</sup>S amino acid content percentages are calculated for proteins in the unprocessed state (containing putative signal peptides).

<sup>c</sup>The low values observed under this condition reflects the fact that some transcripts that increase during S deprivation also show a slight increase during –P conditions (D. González-Ballester and A.R. Grossman, unpublished data).

required for the acetyl cycle of Arg, are diminished. Accumulation of transcripts required for the biosynthesis of CoA in the *snrk2.1* mutant during –S conditions (Table 6) may be indicative of a CoA deficiency. While elevated levels of *ARG9* and *GSD1* gene products could also contribute to the production of Pro, a plant stress-related metabolite, this seems unlikely since there is essentially no increase in the transcript encoding pyrroline-5-carboxylate reductase (PCR1) (Table 6, Figure 11).

Putrescine and other polyamines can also be synthesized via agmatine (Figure 11), although only one transcript of this pathway was elevated in S-deprived *snrk2.1* cells (the agmatine iminohydrolase, *AIH1*), and its absolute level was much lower than that of *ODC2* (Table 6). Hence, as proposed previously (Voigt et al., 2000), the most prominent pathway for putrescine biosynthesis in *Chlamydomonas* is likely via Orn.

The above observations suggest a greater capacity for putrescine synthesis in S-deprived *snrk2.1* than in wild-type cells. Accumulation of putrescine was previously observed in *Arabidopsis* plants deprived of S (Hirai et al., 2003; Nikiforova et al., 2005b). A SAM deficiency was proposed to be the reason for this accumulation since the synthesis of the putrescine-derived polyamines spermidine and spermine requires SAM. In *Chlamydomonas* wild-type cells, transcripts encoding enzymes for spermine and spermidine synthesis (*S*-adenosylmethionine decarboxylase, *DCA1*; spermidine synthase, *SPD1*; and spermine synthase, *SPS1*) declined following S deprivation. This decline was not observed in the *snrk2.1* mutant (Table 6, Figure 11), although spermidine and spermine biosynthesis in the S-starved mutant cells may still be very low as these cells are likely to be deficient for SAM. While the potential accumulation of putrescine in S-deprived *snrk2.1* mutant cells may be a consequence of reduced SAM levels, this accumulation may also be important for acclimation of the organism to its extreme environment. Some have suggested roles for polyamines in the stabilization of membranes, scavenging of ROS, retardation of senescence, and hypersensitive cell death (Yoda et al., 2003; Groppa and Benavides, 2008). Polyamines have also been found to associate with LHCb proteins, influencing nonphotochemical quenching,

photophosphorylation, and organization of PSII. Moreover, treatment of plants with UV-B irradiation and high light can cause an increase in the levels of polyamines associated with LHCb (Del Duca et al., 1995; Della Mea et al., 2004; Ioannidis et al., 2006; Ioannidis and Kotzabasis, 2007; Navakoudis et al., 2007; Sfichi-Duke et al., 2008). Overall, both our work (presented here) and the work of others suggest that oxidative stress, cell death, PSII architecture, and putrescine accumulation may be processes that are closely related.

### Effects of the *snrk2.1* Lesion on Transcript Accumulation during Nutrient-Replete Conditions

Some transcripts exhibited aberrant levels in the *snrk2.1* mutant relative to wild-type cells even when the cells were maintained in nutrient-replete conditions (see Supplemental Data Sets 2D and 2E online). The high number of transcripts in this group (89 in total) suggests that regulation via SNRK2.1 is not simply an on-off switch, but may be more similar to a rheostat that controls cellular responses over a dynamic range of S concentrations. Transcripts elevated in *snrk2.1* relative to wild-type cells maintained in either +S or –S conditions include those encoding a pyruvate ferredoxin oxidoreductase (PFR1), a specific apoferritin (FDX5), an iron hydrogenase (HYD1), the hydrogenase assembly factors HYDEF and HYDG, and several prolyl 4-hydroxylases (Table 6; see Supplemental Data Sets 2C1 and 2D online). The protein products of these genes have previously been associated with cells experiencing anoxia and H<sub>2</sub> production (Mus et al., 2007). The link between SNRK2.1 kinase and repression of genes regulated by anoxic conditions is intriguing but could reflect an indirect impact of S deprivation on the physiological state of the cells. For example, a decrease in photosynthetic activity and an increase in *HYD* transcript accumulation and H<sub>2</sub> production was observed in *Chlamydomonas* transformants harboring antisense constructs of the chloroplast SO<sub>4</sub><sup>2-</sup> permease (SULP1) grown in nutrient-replete medium (Chen et al., 2005). The authors suggested that the reduced permease activity in the transformants limited the supply of S to

**Table 6.** Expression Levels of Selected Genes That Show Aberrant Behavior in the *snrk2.1* Mutant

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM)			
		Wild Type -S/ Wild Type +S		<i>snrk2.1</i> / Wild Type (+S)		<i>snrk2.1</i> / Wild Type (-S)		Wild	Wild	<i>snrk2.1</i>	<i>snrk2.1</i>
		S	Q	S	Q	S	Q	Type-S	Type+S	-S	+S
<b>Redox processes and singlet oxygen stress-related genes</b>											
142363	<i>PRX1</i> , 2-cys peroxiredoxin, thioredoxin dependent peroxidase (chl.)	-2.2		-0.2		-3.1		177	843	20	759
157173	<i>PRX5</i> , peroxiredoxin type II, thioredoxin dependent peroxidase (chl.)	-0.2		-0.6		-1.5		105	124	38	83
186597	<i>APX1</i> , L-ascorbate peroxidase (chl.)	0.3		-0.2		-2.4		103	85	20	74
165193	<i>APX3</i> , L-ascorbate peroxidase (chl.)	-1.4		-0.6		-1.7		41	108	13	70
182933	<i>FSD1</i> , superoxide dismutase [Fe] (chl.)	-0.5		0.3		-1.7		171	241	53	293
53941	<i>MSD1</i> , superoxide dismutase (mito.)	0.0		0.0		-1.7		381	393	120	391
158519	Singlet oxygen stress-related gene	0.8		1.3		-2.4		259	149	48	355
143122	<i>GPXH</i> (GPX5), glutathione peroxidase (cyt.)	0.7		-0.2		2.0		125	78	502	68
193661	<i>GSTS1</i> , glutathione S-transferase	0.1		-0.3		3.7		37	34	494	28
195887	<i>TRXh1</i> , thioredoxin (cyt.)	-0.2		0.0		1.8		190	216	643	210
196117	<i>TRXh2a</i> , thioredoxin (cyt.)	0.3		0.8		2.7		60	49	385	86
196116	<i>TRXh2b</i> , thioredoxin (cyt.)	0.3		0.8		2.7		71	57	455	102
196118	<i>TRXh2c</i> , thioredoxin (cyt.)	0.3		0.8		2.7		75	61	483	108
134747	<i>TRXy</i> , thioredoxin (chl.)	-0.3		0.2		1.9		16	21	60	23
182094	<i>NRX2</i> , nucleoredoxin	-1.7		1.5		7.2		5	15	701	42
141568	<i>NRX3</i> , nucleothioredoxin	-0.8		2.3		6.6		2	3	150	12
195611	<i>GRX2</i> , glutaredoxin, CPYC type (cyt.)	0.5		-0.1		1.8		111	79	384	73
187796	<i>MSRA5</i> , peptide Met sulfoxide reductase	0.4		0.4		1.5		19	14	54	19
139781	<i>TRXf1</i> , thioredoxin f1 (chl.)	0.1		-0.1		-1.8		185	176	52	170
195886	<i>TRXf2</i> , thioredoxin f2 (chl.)	-1.0		-0.4		-2.2		55	111	12	83
<b>Apoptosis-related genes</b>											
141763	<i>DAD1</i> , antiapoptosis gene	-0.6		0.4		-1.7		80	118	24	159
182542	<i>DDI1</i> , DNA damage-inducible v-SNARE binding protein	-0.3		0.2		1.8		9	11	30	12
174932	<i>LAG1</i> , longevity assurance protein	0.0		0.5		3.7		21	20	276	29
205713	Metacaspase, type I	0.4		0.6		2.0		9	6	33	10
205701	<i>PHD1</i> , V1AF phosphatase-like protein	-0.8		0.4		4.5		20	35	437	47
190610	<i>PDCD6/ALG-2</i> , programmed cell death-related gene	0.1		0.6		1.6		23	22	68	33
179155	Senescence-associated gene	-0.1		0.8		4.5		13	14	281	24
<b>Arg/Om/polyamine biosynthesis</b>											
206062	<i>ODC2</i> , Orn decarboxylase 2	0.7		0.2		2.7		45	27	289	30
139007	<i>ARG9</i> , acetylornithine aminotransferase	-0.4		-0.6		2.4		33	45	174	29
183928	<i>DCA1</i> , S-adenosylmethionine decarboxylase	-1.8		-0.3		1.5		54	184	149	146
206050	<i>SPD1</i> , spermidine synthase	-0.9		-0.5		0.7		12	24	20	17
206069	<i>SPS1</i> , spermine synthase	-2.6		-1.0		2.0		6	34	22	17
193221	<i>PRD</i> , Pro dehydrogenase	0.5		1.0		2.8		14	10	100	19
192364	<i>PCR1</i> , pyrroline-5-carboxylate reductase	-1.3		0.2		1.0		15	37	29	42
130812	<i>GSD1</i> , Glu-5-semialdehyde dehydrogenase	-0.6		-0.5		2.0		42	62	164	45
195386	<i>OAT1</i> , Orn aminotransferase	1.6		0.9		-1.2		48	16	20	30
13028	<i>AIH1</i> , agmatine iminohydrolase	-0.2		-1.0		3.5		3	3	28	2
<b>Hydrogen production-related genes</b>											
156833	<i>FDX5</i> , Apoferreredoxin	-2.6		6.2		7.7		1	5	174	367
122198	<i>PFR1</i> , pyruvate-ferreredoxin oxidoreductase	1.8		6.7		6.5		1	0	103	33
183963	<i>HYD1</i> , iron hydrogenase	0.4		1.9		2.0		49	36	190	130
128256	<i>HYDEF</i> , iron hydrogenase assembly protein	0.8		2.0		1.6		12	7	35	26
196226	<i>HYDG</i> , hydrogenase assembly factor	1.0		1.7		1.4		61	30	158	97
<b>Carbon metabolism</b>											
115491	Acetate transporter (p)	-1.7		-1.5		-1.6		699	2285	228	806
196311	<i>ACS3</i> , Acetyl CoA synthetase, mit	-1.6		-1.4		-1.1		233	710	112	271

(Continued)

**Table 6.** (continued).

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM)			
		Wild Type –S/ Wild Type +S		<i>snrk2.1</i> / Wild Type (+S)		<i>snrk2.1</i> / Wild Type (–S)		Wild Type–S	Wild Type+S	<i>snrk2.1</i> –S	<i>snrk2.1</i> +S
		S	Q	S	Q	S	Q				
24120	<i>CAH1</i> , carbonic anhydrase (ext.)	–2.1	–4.1	1.9	6.4	2.4	8.2	328	1375	1727	5057
24552	<i>CAH4</i> , carbonic anhydrase	–2.2	0.6	2.1	8.3	3.0	6.1	83	368	655	1614
55019	<i>LC11</i> , low CO <sub>2</sub> -inducible gene	–2.9		2.5		3.3		20	150	199	843
189430	<i>CCP1</i> , chl. envelope protein (chl.)	–0.7		1.9		1.9		139	230	529	830
128766	<i>CCP2</i> , chl. envelope protein (chl.)		0.3	2.2		2.3		90	76	453	346
135648	<i>NAR1.2</i> , bicarbonate transporter (p)	–0.7		2.0		1.6		132	214	401	849
205996	Limiting CO <sub>2</sub> -inducible gene	–0.4		1.9		2.5		25	33	139	124
CoA biosynthesis											
116788	Phosphopantothenate Cys ligase	0.0		–0.7		4.1		6	6	103	4
196344	PANB, ketopantoate hydroxymethyltransferase	–0.7		0.4		4.3		3	5	56	6

Details are as given in Table 3.

chloroplasts, which caused a decline in PSII activity and a decrease in O<sub>2</sub> evolution. This situation could result in a rate of respiratory O<sub>2</sub> uptake that exceeds the rate of O<sub>2</sub> evolution, and the cells would experience anaerobic conditions even in the light growing on complete medium. The *snrk2.1* mutant cultures may also be experiencing a low level of S limitation during nutrient-replete growth since the levels of many transcripts encoding proteins required for efficient uptake and assimilation of SO<sub>4</sub><sup>2–</sup> were significantly reduced in the mutant relative to wild-type cells (Table 3). Thus, *snrk2.1* may have a greater tendency to become anoxic and produce H<sub>2</sub> than wild-type cells during nutrient-replete growth.

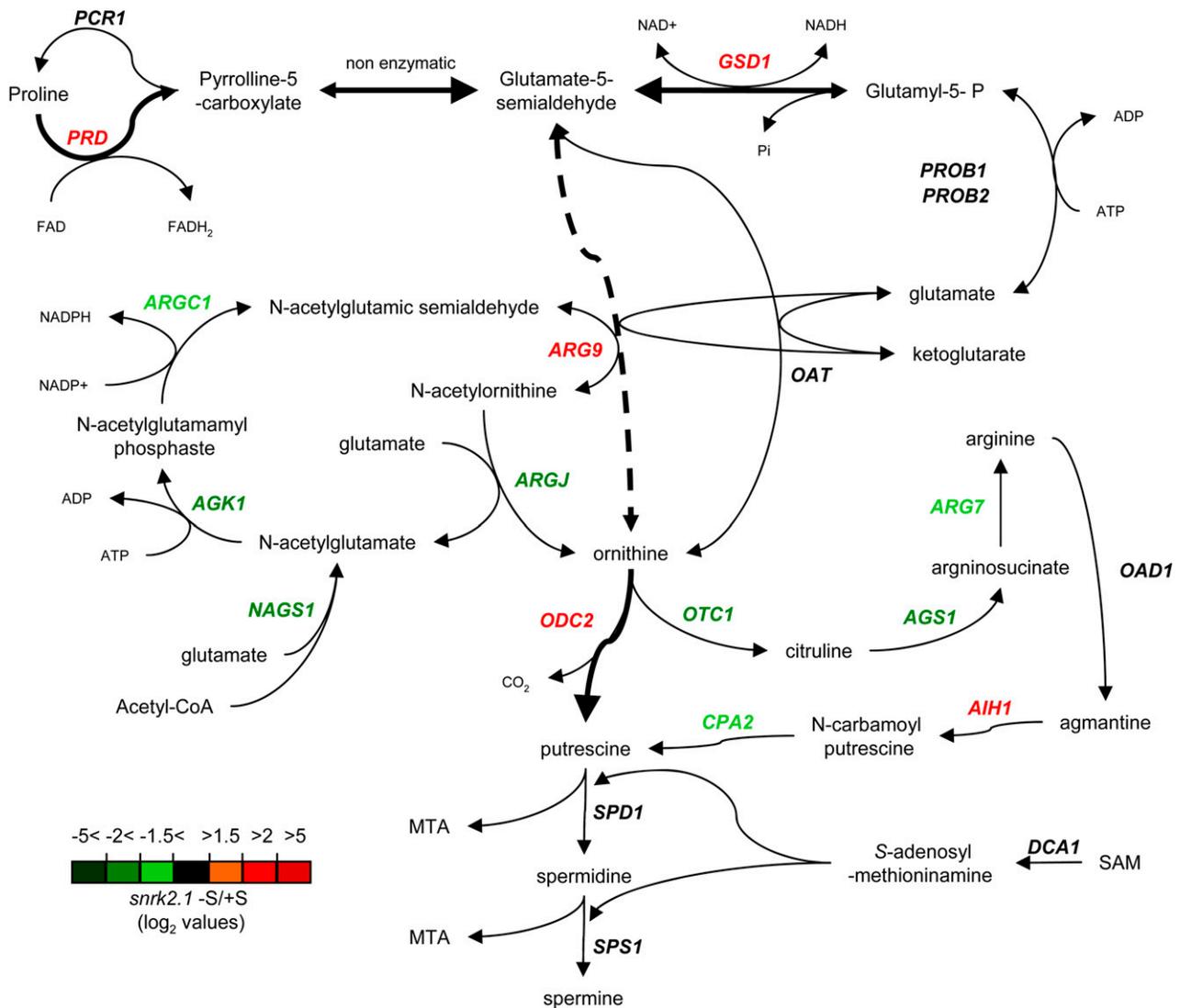
Interestingly, even when *snrk2.1* cells were maintained in acetate-containing, nutrient-replete medium, which normally would partially suppress low CO<sub>2</sub> responses, many CCM-related transcripts became highly abundant. These transcripts include those encoding the periplasmic carbonic anhydrase *CAH1*, mitochondrial *CAH4*, the HCO<sub>3</sub><sup>–</sup>/nitrite transporter *NAR1.2* (Mariscal et al., 2006), and several previously identified low CO<sub>2</sub>-inducible genes (see Supplemental Data Sets 2C1 and 2D online). The *CAH1* transcript level was strongly repressed in the *snrk2.1* mutant in acetate-free medium bubbled with 4% CO<sub>2</sub> in air (similar to wild-type cells), confirming that the *snrk2.1* lesion has an indirect effect on the regulation of this gene and potentially other CCM genes in acetate-containing medium (see Supplemental Figure 3 online). Interestingly, the *snrk2.1* mutant grown in nutrient-replete medium exhibited low levels of transcripts (relative to wild-type cells) encoding proteins required for acetate assimilation, including a putative acetate transporter (ID 115491) and the acetyl CoA synthetase (*ACS3*) (Table 6). These data suggest that the *snrk2.1* strain may favor the use of atmospheric CO<sub>2</sub>/HCO<sub>3</sub><sup>–</sup> even when growing in acetate-containing medium. Normally, *Chlamydomonas* cells grown in acetate-containing medium exhibit decreased rates of CO<sub>2</sub> fixation and O<sub>2</sub> evolution (Heifetz et al., 2000). This bias of *snrk2.1* cells toward the fixation of CO<sub>2</sub> rather than toward the assimilation of acetate may in some way link to the greater tendency of the mutant culture to become anoxic. Elevated CO<sub>2</sub> fixation would

potentially increase O<sub>2</sub> evolution and reduce the level of anoxia in mutant cultures grown in the light on nutrient-replete, acetate-containing medium. These relationships and their physiological implications need to be more thoroughly established.

## Conclusion

In this study, the use of RNA-seq technology yielded deep coverage for most individual transcripts in *Chlamydomonas* and provided a genome-scale, quantitative view of how a changing S environment impacts transcript patterns. Analysis of the transcriptome profiles of both the wild type and the *snrk2.1* mutant, which is unable to acclimate to S deprivation, clearly identified the set of –S-responsive genes, and those –S-responsive genes under the control of the regulatory SNRK2.1 kinase. S-deprived *Chlamydomonas* cells accumulated transcripts involved in the scavenging of environmental S and in the reductive assimilation of SO<sub>4</sub><sup>2–</sup> to Cys, while transcripts encoding enzymes required for the synthesis of Met, SAM, thiamine, and biotin declined. This result suggests the importance of maintaining an efficient capacity for Cys biosynthesis relative to other S-containing metabolites. Moreover, the results of this study also indicate a massive redistribution of the S resources of the cells as exemplified by the coordinated S deprivation-elicited accumulation of the *CDO1*, *TAUD1*, *TAUD2*, *RDP3*, *CSD2*, and *SOUX1* transcripts. While the enzymes encoded by these transcripts have been associated with the redistribution of S through the catabolism of Cys and other S-containing molecules (e.g., sulfolipids), their exact substrate specificities have not been defined. Indeed, the precise ways that these enzymes participate in creating a new S homeostasis during S deprivation have not yet been established for any photosynthetic organism.

The critical nature of the redistribution of S resources is also suggested by changes in the organization of cellular processes during S deprivation. A massive increase in the transcript encoding LHCBM9, a light-harvesting protein with a reduced S amino acid content relative to other LHCBM polypeptides,



**Figure 11.** Potential Pathways for Polyamine Biosynthesis.

Genes encoding the proteins of the pathways are labeled with colors, indicating the fold change ( $\log_2$ ) of their transcript levels under S deprivation conditions relative to nutrient-replete conditions for *snrk2.1* cells based on RNA-seq data (color code is given in the figure). The dashed line indicates the alternative activities proposed for the ARG9 protein. The protein encoded by each of the genes represented in the pathway diagram is given in Table 6. MTA, S-methyl-5'-thioadenosine.

paralleled a decline in the level of transcripts encoding all of the other LHCBM9s. This shift in antenna protein production could lead to significant S conservation and a change in the ways in which the alga uses light energy. Indeed, at a global level, proteins that accumulate during  $-S$  conditions tend to have a reduced S amino acid content; such proteins include the ECPs, HAPs, ARSs, and TAUDs. Based on these findings, it is likely that a recycling of S-containing amino acids and their reallocation among proteins with low S amino acid content represents an extreme economy of S use that can extend cell viability when S resources become scarce.

Another important aspect of this work addresses the question of what a cell does when it is unable to implement the  $-S$

acclimation program, as is the case of the *snrk2.1* mutant. The inability of mutant cells to acclimate to  $-S$  conditions likely prevents efficient use of external  $SO_4^{2-}$  (free and organic), the recycling of internal S (SQDG and Cys), and the redistribution of the S resources among different S-containing metabolites (e.g., CoA and SAM); this in turn results in aberrant control of cellular metabolism. Moreover, the absence of transcripts encoding LHCBM9 and LHC9SR1 in the *snrk2.1* mutant during  $-S$  conditions may alter the structure of the photosynthetic apparatus. These metabolic and structural aberrations can result in the accumulation of metabolites and reactive molecules (e.g., putrescine and  $^1O_2$ ) that have extreme consequences, ultimately leading to diminished cell viability. Indeed, when the *snrk2.1*

mutant is starved for S, it appears to show molecular responses, such as the accumulation of transcripts associated with oxidative damage, ROS production, and apoptosis that are typical of organisms experiencing extreme environmental conditions.

## METHODS

### Strains and Culture Conditions

*Chlamydomonas reinhardtii* strains used for this study were D66 (wild-type strain; *nit2 cw15 mt+*) (Pollock et al., 2003), *ars11 (snrk2.1cw15mt+)* (González-Ballester et al., 2008), 21gr (wild-type strain; *nit5 mt-*) (Harris, 1989), and *sac1 (sac1mt+)* (Davies et al., 1994). The *ars11* strain was designated as the *snrk2.1* mutant throughout since the lesion is in the *SNRK2.1* gene. Cells were cultured under continuous light of  $\sim 60 \mu\text{mol photon m}^{-2} \text{ s}^{-1}$  at 23°C in liquid and on solid Tris-acetate-phosphate (TAP) medium (Harris, 1989). To impose S deprivation, cells in mid-logarithmic growth phase were washed twice with liquid TAP medium without S (TAP-S) (Harris, 1989), and equal numbers of cells were resuspended in TAP or TAP-S. Cell aliquots were collected for RNA isolation just prior to and 1, 6, and 24 h after being transferred to TAP and TAP-S medium.

### RNA Isolation, Synthesis of cDNA, and qRT-PCR

Isolation of total RNA using a phenol-chloroform procedure, cDNA synthesis, and qRT-PCR was performed by standard methods as previously described (Schloss et al., 1984; González-Ballester et al., 2008). Ct values were determined for three independent technical experiments, with at least two measurements of the same sample in each experiment. Relative fold differences were calculated based on the  $\Delta\text{Ct}$  method ( $2^{-(\text{Ct}_{\text{sample}} - \text{Ct}_{\text{control gene}})}$ ) (Livak and Schmittgen, 2001) using the *CBLP* amplification product as an internal standard. Primer pairs used for qRT-PCR analyses are in Supplemental Table 2 online. Sizes of amplification products were 100 to 300 bp.

### Microarray Analysis

Oligonucleotides ( $\sim 70$  nucleotides length) representing  $\sim 10,000$  unique ESTs were used to print the arrays. The design and generation of the microarrays chips used in this work was described previously (Eberhard et al., 2006). The following RNA samples were compared in the microarrays: wild type 1 h -S compared with wild type 1 h +S, wild type 6 h -S compared with wild type 6 h +S, wild type 24 h -S compared with wild type 24 h +S, wild type 0 h -S compared with *snrk2.1* 0 h -S, wild type 1 h -S compared with *snrk2.1* 1 h -S, wild type 6 h -S compared with *snrk2.1* 6 h -S, and wild type 24 h -S compared with *snrk2.1* 24 h -S. Preparation of the fluorescent microarray probes and hybridization to the arrays was performed as described previously (Eberhard et al., 2006; Mus et al., 2007). Between two and four biological replicates were used for each sample, and at least four microarray slides for each comparison (additionally, each slide has two copies of each array element). Arrays were scanned using a GenePix 4000B scanner (Molecular Devices). Photomultiplier tube voltages were adjusted manually to balance Cy5 and Cy3 signals (1:1 ratio) and to avoid spot saturation. Resulting images were imported into SpotReader (version 1.3.0.5; Niles Scientific), and spot positions were defined and fluorescence signal intensities determined. Spot quality signals were automatically flagged as reliable (good), non-reliable (absent), or empty based on the following parameters: saturated foreground, 70%; uneven color, 0.5; bright specks, 4; low foreground, 1.4; high background, 10; low background, 0.1; variable background, 5. Additionally, spot signals that were distorted by dust, locally high backgrounds, or printing flaws were manually flagged as nonreliable (absent). Spot backgrounds were automatically subtracted; a radius of 2 times the

average radius for each spot (from the spot center) was defined, and the region of the field within this radius and between the new diameter and the original spot diameter was considered background. Subsequent data analyses were performed using GeneSpring 7.3.1 (Agilent Technologies). Normalizations of spot intensities were achieved using the software standard "per spot and per chip intensity-dependent (Lowess)" normalization (<http://stat-www.berkeley.edu/users/terry/zarray/Html/normspie.html>). Only spots flagged as reliable (good) were analyzed. Error models were computed based on replicates. Signal ratios were considered only if they passed Student's *t* test for significance with a P value of  $\leq 0.05$ .

Microarray probe sequences were mapped to the *Chlamydomonas* v.3.1 genome to find their correspondence with gene models. Gapped alignment using BLAT (Kent, 2002) was performed for a set of 9998 microarray probes, of which 6745 overlapped with JGI v.3.1 gene models. At least one matching probe was found for 5948 gene models (out of 14,598). When genes have multiple matching probes, the probe with a more significant fold change was selected when comparing two samples. Correspondences between microarray probes and JGI v.3.1 gene models are included in Supplemental Data Set 1 online.

### RNA-Seq Technology

#### Sample Preparation

Total RNA from wild-type (D66) and *snrk2.1* mutant cells 6 h after being transferred to -S and +S medium were submitted to Illumina for sequencing using their proprietary Genome Analyzer. cDNA libraries were assembled according to the manufacturer's RNA-seq protocol, loaded, and sequenced as 35-mers in a total of 11 lanes of the sequencer (see Supplemental Table 3 online). Raw image files were collected by the sequencer and analyzed using the standard Illumina pipeline.

#### Mapping Reads to the Genome

Processed sequence files from the Illumina pipeline output were aligned against the v.3.1 assembly of the *Chlamydomonas* genome. The SOAP alignment program (Li et al., 2008) was used to map the short reads in two steps (see Supplemental Methods online for details). In a first round, the raw 35-mers sequences were aligned to the genomic sequence with a tolerance of up to two mismatches and no indels. Approximately 77% of the reads mapped to at least one genomic location on either of the two strands, and of these, 89% mapped uniquely. This percentage compares favorably with previously reported RNA-seq studies (Marioni et al., 2008; Mortazavi et al., 2008; Nagalakshmi et al., 2008). For those sequences that did not align to the genome, a second round alignment was performed in which we recursively trimmed one base at a time (up to 21-mers) at either the 5' or 3' end of the reads until we found a (perfect) match to the genome. This procedure allowed us to recover some misalignments caused by the presence of exon-exon junctions, sequencing errors, adaptor sequences, and poly(A) tails in the sequence reads. About 50% of the reads that were not aligned in the initial analysis were recovered, bringing the total fraction of mappable reads to  $\sim 88\%$  (see Supplemental Table 3 online).

All alignments were tagged as either unique or not unique and then combined. The most 5' position of each alignment was recorded along with the length *k* of the aligned read (*k-hit* throughout). For differential expression analysis, per-base unique *k-hits* from both strands were pooled (as the cDNA library lacks strand specificity) and summed across the genome to obtain gene counts.

#### Transcript Abundance

Relative transcript abundances were determined as follows (see Supplemental Methods online for notation): for each particular gene model *j* in

the JGI annotation and each read length  $k$  (in the range 21 to 35 in this work), we determined the set of unambiguously mappable bases ( $M_{jk}^m$ ) at a given multiplicity threshold  $m$ , and its cardinality  $M_{jk}^m = |M_{jk}^m|$  denotes the number of mappable bases per gene or  $m$ -mappability. The mappable set  $M_{jk}^1$  includes all of the exonic bases that are uniquely mappable with  $k$ -mers, except for those that overlap with neighbor gene models or are located in a window of size  $k - 1$  at the 3' end of the exon. For each lane  $i$ , we quantified the number of unique  $k$  hits associated with each gene ( $H_{ijk}^1$ ), normalized the value using the corresponding  $1$ -mappability, and summed over all  $k$  values using the total number of aligned sequences  $H_i$  (a measure of sequencing depth) as a global scaling factor. The transcript relative abundance  $T_{ij}^1$  (TRA) is expressed in units of reads per kilobase of mappable exon length per million mapped reads (RPKM), where the constant factor  $c$  is equal to  $1 \times 10^6$  (Mortazavi et al., 2008). More details are shown online in Supplemental Methods, Supplemental Figure 6, and Supplemental Figure 7.

$$T_{ij}^1 = \frac{c}{H_i} \sum_k \frac{H_{ijk}^1}{M_{jk}^1}$$

### Reproducibility of the Data

The total number of hits  $H_{ij}^1$  for a particular gene was modeled as a random variable from a Poisson distribution. The significance of hits-per-gene variation between different lanes, runs, or experiments was calculated as described previously (Audic and Claverie, 1997) (see Supplemental Methods online for details): the conditional probability,  $p(H_{a(ij)}^1 | H_{b(ij)}^1)$  for the number of hits in two different groups of lanes ( $a(i)$ ,  $b(i)$ ) for the same gene model  $j$  is computed by means of a uniform prior for all values of the parameter of the Poisson distribution. Thus, the cumulative distribution function of  $p(H_{a(ij)}^1 | H_{b(ij)}^1)$  is used to estimate the probability (designated P value throughout) that the observed hits fall outside the confidence interval defined by  $H_{a(ij)}^1$  and  $H_{b(ij)}^1$ .

Based on the previous statistical framework, pairwise comparisons were performed between and within runs and/or samples to get a sense of the quality and reproducibility of the sequencing process. The null hypothesis (that the number of hits from both lanes follows the same random distribution) can be tested from deviations of the P value distribution from uniformity. Supplemental Figure 4 online shows qq plots (quantile probability plots) of hits-per-gene statistics for several pairs of lanes. Supplemental Figure 5 online shows normalized expression fold changes relative to overall expression. The data shows that lane-to-lane or run-to-run comparisons show slight overrepresentations of small P values. However, sample-to-sample variations are much more pronounced. Additional details regarding the false discovery rate for differential expression can be found in Supplemental Methods online.

### Estimation of Differential Expression

Due to small but detectable lane and flow cell effects, we chose to take the mean ( $\bar{T}_{a(ij)}^1$ ) of single-lane TRA values for each condition as if they were true experimental replicates to average out the impact of sequencing bias in expression estimates. By contrast, P values were computed after pooling all lanes for the same sample to improve detection of low abundance transcripts and minimize the effect of different levels of sequencing depth across samples. Moreover, this choice also allows us to make use of true count data using the aforementioned Poisson approach.

Fold changes in expression between conditions (wild type –S/wild type +S, *snrk2.1* –S/*snrk2.1* +S, *snrk2.1* –S/wild type –S, and *snrk2.1* +S/wild type +S) were calculated using the TRAs  $\log_2$  ratios. For those gene models with null TRA values, a value equal to the minimum on each experiment was imputed. In that case, fold changes must be interpreted as lower bounds of the actual value. The sets of genes that were

considered for further analysis in each comparison were selected after applying the following filters: (1) TRA  $\log_2$  ratios were considered significant if  $\geq 1.5$  or  $\leq -1.5$ ; (2) positive TRA  $\log_2$  ratios that had numerators below 30 were not considered; (3) negative TRA  $\log_2$  ratios that had denominators below 30 were not considered; (4) the P value for differential expression was set to be  $\leq 0.05$ . After applying the three first filters (1 to 3), >99.9% of the genes passed the fourth filter. Transcripts that passed all the filters were used to derive the information in Figure 3 and Supplemental Data Set 2 online.

### Sequence Analysis

Gene annotations were retrieved from the *Chlamydomonas* genome browser, which was developed by the JGI (<http://genome.jgi-psf.org/Chlre3/Chlre3.home.html>) in collaboration with the *Chlamydomonas* community (Merchant et al., 2007). For gene models with no associated annotation, BLAST searches were performed to identify sequence similarities. Sequences were analyzed using DNASTAR software v.4.05 (Lasergene Navigator), Bioedit Sequence Alignment Editor version 5.0.9 (Department of Microbiology, North Carolina State University), the National Center for Biotechnology Information BLAST server (<http://www.ncbi.nlm.nih.gov/BLAST/>), and the iPSORT and PSORT servers (<http://hc.ims.u-tokyo.ac.jp/iPSORT/> and <http://psort.ims.u-tokyo.ac.jp/form.html>).

### Accession Numbers

JGI v.3.1 protein accession numbers are provided in Supplemental Data Set 1 online with the corresponding Web link to the JGI browser. RNA-seq raw data (see Supplemental Data Sets 1 to 3 online) are available at the Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo/>) under accession number GSE17970 and at <http://genomes.mcdb.ucla.edu/CreSulfur/>.

### Supplemental Data

**Supplemental Figure 1.** *Chlamydomonas* Growth with Different S-Containing Amino Acids as Sole S Source.

**Supplemental Figure 2.** Alignment of Putative Taurine Dioxygenases (TAUD).

**Supplemental Figure 3.** *CAH1* Transcript Accumulation in Wild-Type and *snrk2.1* Cells under Fixed and Inorganic Carbon Conditions (Based on qRT-PCR).

**Supplemental Figure 4.** qq-Plots Comparing P Values from Selected Pairwise Comparisons of the Total Number of Hits per Gene from RNA-Seq Individual Lanes (See Supplemental Table 3 Online for Details).

**Supplemental Figure 5.** Logarithmic Scatterplots of Normalized Expression Fold Changes ( $y$  Axes) versus Overall Expression Measures ( $x$  Axes).

**Supplemental Figure 6.** Definition of Mappable Sets and Assignment of  $k$ -Hits per Gene.

**Supplemental Figure 7.** Coverage Data for a 3-kb Region of the *GAP3* Locus.

**Supplemental Table 1.** *Chlamydomonas* Singlet Oxygen Stress-Related Genes and Their Expression in the Wild Type and *snrk2.1* Mutant in –S Conditions.

**Supplemental Table 2.** Primers Used for qPCR Analyses.

**Supplemental Table 3.** Summary Statistics for RNA-Seq Samples and Sequence Alignment.

**Supplemental Data Set 1.** Expression Levels of All Analyzed Transcripts.

**Supplemental Data Set 2.** Relative Levels of Transcripts Represented in Figure 3.

**Supplemental Data Set 3.** Expression Levels of Genes Grouped by Putative Functions.

**Supplemental Methods.**

**Supplemental References.**

## ACKNOWLEDGMENTS

We thank Jeffrey Moseley and Wirulda Pootakham for critical discussions and for reading the manuscript and the Joint Genome Institute for generating a *Chlamydomonas* genome sequence, which has been invaluable for both gene identification and mutant analyses. The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We also thank Weihong Yan for loading the data onto the genome browser. This work was supported by National Science Foundation Grants MCB 0824469 and MCB 0235878 awarded to A.R.G., as well as funds provided by the Carnegie Institution, by the Marie Curie MOIF-CT-2006-40208-APOSD grant (EU) awarded to D.G.-B., and by the Spanish Foundation of Science and Technology (FECYT) and the California Nanosystems Institute (UCLA) grants awarded to D.C. S.S.M. and M.P. acknowledge support from the Department of Energy (DE-FC02-03ER63421).

Received September 4, 2009; revised April 1, 2010; accepted May 18, 2010; published June 29, 2010.

## REFERENCES

- Apel, K., and Hirt, H. (2004). Reactive oxygen species: Metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* **55**: 373–399.
- Asada, K. (2006). Production and scavenging of reactive oxygen species in chloroplasts and their functions. *Plant Physiol.* **141**: 391–396.
- Audic, S., and Claverie, J.M. (1997). The significance of digital gene expression profiles. *Genome Res.* **7**: 986–995.
- Bentley, R., and Chasteen, T.G. (2004). Environmental VOSCs-formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere* **55**: 291–317.
- Billheimer, J.T., Cornvale, H.N., Leisinger, T., Eckhardt, T., and Jones, E.E. (1976). Ornithine  $\delta$ -transaminase in *Escherichia coli*: Its identity with acetylornithine  $\delta$ -transaminase. *J. Bacteriol.* **127**: 1315–1323.
- Bolling, C., and Fiehn, O. (2005). Metabolite profiling of *Chlamydomonas reinhardtii* under nutrient deprivation. *Plant Physiol.* **139**: 1995–2005.
- Bordo, D., and Bork, P. (2002). The rhodanese/Cdc25 phosphatase superfamily. Sequence-structure-function relations. *EMBO Rep.* **3**: 741–746.
- Chang, C.W., Moseley, J.L., Wykoff, D., and Grossman, A.R. (2005). The *LPB1* gene is important for acclimation of *Chlamydomonas reinhardtii* to phosphorus and sulfur deprivation. *Plant Physiol.* **138**: 319–329.
- Chen, H.C., Newton, A.J., and Melis, A. (2005). Role of SulP, a nuclear-encoded chloroplast sulfate permease, in sulfate transport and H<sub>2</sub> evolution in *Chlamydomonas reinhardtii*. *Photosynth. Res.* **84**: 289–296.
- Collier, J.L., and Grossman, A.R. (1992). Chlorosis induced by nutrient deprivation in *Synechococcus* sp. Strain PCC 7942: Not all bleaching is the same. *J. Bacteriol.* **174**: 4718–4726.
- Colnaghi, R., Cassinelli, G., Drummond, M., Forlani, F., and Pagani, S. (2001). Properties of the *Escherichia coli* rhodanese-like protein SseA: Contribution of the active-site residue Ser240 to sulfur donor recognition. *FEBS Lett.* **500**: 153–156.
- Couee, I., Sulmon, C., Gouesbet, G., and El Amrani, A. (2006). Involvement of soluble sugars in reactive oxygen species balance and responses to oxidative stress in plants. *J. Exp. Bot.* **57**: 449–459.
- Davies, J., Yildiz, F., and Grossman, A.R. (1996). *Sac1*, a putative regulator that is critical for survival of *Chlamydomonas reinhardtii* during sulfur deprivation. *EMBO J.* **15**: 2150–2159.
- Davies, J.P., Yildiz, F., and Grossman, A.R. (1994). Mutants of *Chlamydomonas reinhardtii* with aberrant responses to sulfur deprivation. *Plant Cell* **6**: 53–63.
- Davies, J.P., Yildiz, F.H., and Grossman, A.R. (1999). *Sac3*, an Snf1-like serine/threonine kinase that positively and negatively regulates the responses of *Chlamydomonas* to sulfur limitation. *Plant Cell* **11**: 1179–1190.
- Dayer, R., Fischer, B.B., Eggen, R.I., and Lemaire, S.D. (2008). The peroxiredoxin and glutathione peroxidase families in *Chlamydomonas reinhardtii*. *Genetics* **179**: 41–57.
- de Hostos, E.L., Schilling, J., and Grossman, A.R. (1989). Structure and expression of the gene encoding the periplasmic arylsulfatase of *Chlamydomonas reinhardtii*. *Mol. Gen. Genet.* **218**: 229–239.
- de Hostos, E.L., Togasaki, R.K., and Grossman, A.R. (1988). Purification and biosynthesis of a derepressible periplasmic arylsulfatase from *Chlamydomonas reinhardtii*. *J. Cell Biol.* **106**: 29–37.
- Del Duca, S., Beninati, S., and Serafini-Fracassini, D. (1995). Polyamines in chloroplasts: Identification of their glutamyl and acetyl derivatives. *Biochem. J.* **305**: 233–237.
- Della Mea, M., Di Sandro, A., Dondini, L., Del Duca, S., Vantini, F., Bergamini, C., Bassi, R., and Serafini-Fracassini, D. (2004). A Zea mays 39-kDa thylakoid transglutaminase catalyses the modification by polyamines of light-harvesting complex II in a light-dependent way. *Planta* **219**: 754–764.
- Dembitsky, V.M. (2003). Oxidation, epoxidation and sulfoxidation reactions catalyzed by haloperoxidases. *Tetrahedron* **59**: 4701–4720.
- Dominy, J.E., Jr., Simmons, C.R., Karplus, P.A., Gehring, A.M., and Stipanuk, M.H. (2006). Identification and characterization of bacterial cysteine dioxygenases: A new route of cysteine degradation for eubacteria. *J. Bacteriol.* **188**: 5561–5569.
- Eberhard, S., Jain, M., Im, C.S., Pollock, S., Shrager, J., Lin, Y., Peek, A.S., and Grossman, A.R. (2006). Generation of an oligonucleotide array for analysis of gene expression in *Chlamydomonas reinhardtii*. *Curr. Genet.* **49**: 106–124.
- Eichhorn, E., van der Ploeg, J.R., Kertesz, M.A., and Leisinger, T. (1997). Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J. Biol. Chem.* **272**: 23031–23036.
- Eirad, D., and Grossman, A.R. (2004). A genome's-eye view of the light-harvesting polypeptides of *Chlamydomonas reinhardtii*. *Curr. Genet.* **45**: 61–75.
- Fauchon, M., Lagniel, G., Aude, J.C., Lombardia, L., Soularue, P., Petat, C., Marguerie, G., Sentenac, A., Werner, M., and Labarre, J. (2002). Sulfur sparing in the yeast proteome in response to sulfur demand. *Mol. Cell* **9**: 713–723.
- Ghirardi, M.L., Posewitz, M.C., Maness, P.C., Dubini, A., Yu, J., and Seibert, M. (2007). Hydrogenases and hydrogen photoproduction in oxygenic photosynthetic organisms. *Annu. Rev. Plant Biol.* **58**: 71–91.

- Giovanelli, J., Mudd, S.H., and Datko, A.H. (1985). Quantitative analysis of pathways of methionine metabolism and their regulation in *Lemna*. *Plant Physiol.* **78**: 555–560.
- González-Ballester, D., and Grossman, A.R. (2009). Sulfur: From acquisition to assimilation. In *The Chlamydomonas Sourcebook*, D. Stern, E.H. Harris, and G.B. Witman, eds (Amsterdam, The Netherlands: Elsevier), pp. 159–188.
- González-Ballester, D., Pollock, S.V., Pootakham, W., and Grossman, A.R. (2008). The central role of a SNRK2 kinase in sulfur deprivation responses. *Plant Physiol.* **147**: 216–227.
- Goyer, A., Haslekas, C., Miginiac-Maslow, M., Klein, U., Le Marechal, P., Jacquot, J.P., and Decottignies, P. (2002). Isolation and characterization of a thioredoxin-dependent peroxidase from *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **269**: 272–282.
- Groppa, M.D., and Benavides, M.P. (2008). Polyamines and abiotic stress: Recent advances. *Amino Acids* **34**: 35–45.
- Hansch, R., Lang, C., Rennenberg, H., and Mendel, R.R. (2007). Significance of plant sulfite oxidase. *Plant Biol.* **9**: 589–595.
- Hanson, A.D., and Roje, S. (2001). One-carbon metabolism in higher plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **52**: 119–137.
- Harris, E.H. (1989). *The Chlamydomonas Sourcebook*. A Comprehensive Guide to Biology and Laboratory Use. (San Diego: Academic Press).
- Heifetz, P.B., Forster, B., Osmond, C.B., Giles, L.J., and Boynton, J.E. (2000). Effects of acetate on facultative autotrophy in *Chlamydomonas reinhardtii* assessed by photosynthetic measurements and stable isotope analyses. *Plant Physiol.* **122**: 1439–1445.
- Heimberg, H., Boyen, A., Crabeel, M., and Glansdorff, N. (1990). *Escherichia coli* and *Saccharomyces cerevisiae* acetylornithine aminotransferase: Evolutionary relationship with ornithine aminotransferase. *Gene* **90**: 69–78.
- Hemrika, W., Renirie, R., Dekker, H.L., Barnett, P., and Wever, R. (1997). From phosphatases to vanadium peroxidases: A similar architecture of the active site. *Proc. Natl. Acad. Sci. USA* **94**: 2145–2149.
- Hirai, M.Y., Fujiwara, T., Awazuhara, M., Kimura, T., Noji, M., and Saito, K. (2003). Global expression profiling of sulfur-starved *Arabidopsis* by DNA microarray reveals the role of O-acetyl-L-serine as a general regulator of gene expression in response to sulfur nutrition. *Plant J.* **33**: 651–663.
- Hogan, D.A., Auchtung, T.A., and Hausinger, R.P. (1999). Cloning and characterization of a sulfonate/alpha-ketoglutarate dioxygenase from *Saccharomyces cerevisiae*. *J. Bacteriol.* **181**: 5876–5879.
- Ioannidis, N.E., and Kotzabasis, K. (2007). Effects of polyamines on the functionality of photosynthetic membrane in vivo and in vitro. *Biochim. Biophys. Acta* **1767**: 1372–1382.
- Ioannidis, N.E., Sfichi, L., and Kotzabasis, K. (2006). Putrescine stimulates chemiosmotic ATP synthesis. *Biochim. Biophys. Acta* **1757**: 821–828.
- Irihimovitch, V., and Stern, D.B. (2006). The sulfur acclimation SAC3 kinase is required for chloroplast transcriptional repression under sulfur limitation in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. USA* **103**: 7911–7916.
- Kahnert, A., Vermeij, P., Wietek, C., James, P., Leisinger, T., and Kertesz, M.A. (2000). The *ssu* locus plays a key role in organosulfur metabolism in *Pseudomonas putida* S-313. *J. Bacteriol.* **182**: 2869–2878.
- Kent, W.J. (2002). BLAT - The BLAST-like alignment tool. *Genome Res.* **12**: 656–664.
- Kessler, D. (2006). Enzymatic activation of sulfur for incorporation into biomolecules in prokaryotes. *FEMS Microbiol. Rev.* **30**: 825–840.
- Kirk, D.L., and Kirk, M.M. (1978). Carrier-mediated uptake of arginine and urea by *Chlamydomonas reinhardtii*. *Plant Physiol.* **61**: 556–560.
- Kopriva, S., Mugford, S.G., Matthewman, C., and Koprivova, A. (2009). Plant sulfate assimilation genes: Redundancy versus specialization. *Plant Cell Rep.* **28**: 1769–1780.
- Ledford, H.K., Chin, B.L., and Niyogi, K.K. (2007). Acclimation to singlet oxygen stress in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **6**: 919–930.
- Ledwidge, R., and Blanchard, J.S. (1999). The dual biosynthetic capability of N-acetylornithine aminotransferase in arginine and lysine biosynthesis. *Biochemistry* **38**: 3019–3024.
- Lemaire, S., Keryer, E., Stein, M., Schepens, I., Issakidis-Bourguet, E., Gérard-Hirne, C., Miginiac-Maslow, M., and Jacquot, J.P. (1999). Heavy-metal regulation of thioredoxin gene expression in *Chlamydomonas reinhardtii*. *Plant Physiol.* **120**: 773–778.
- Lewandowska, M., and Sirko, A. (2008). Recent advances in understanding plant response to sulfur-deficiency stress. *Acta Biochim. Pol.* **55**: 457–471.
- Li, R., Li, Y., Kristiansen, K., and Wang, J. (2008). SOAP: Short oligonucleotide alignment program. *Bioinformatics* **24**: 713–714.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* **25**: 402–408.
- Lunde, C., Zygadlo, A., Simonsen, H.T., Nielsen, P.L., Blennow, A., and Haldrup, A. (2008). Sulfur starvation in rice: The effect on photosynthesis, carbohydrate metabolism, and oxidative stress protective pathways. *Physiol. Plant.* **134**: 508–521.
- Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008). RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**: 1509–1517.
- Mariscal, V., Moulin, P., Orsel, M., Miller, A.J., Fernandez, E., and Galvan, A. (2006). Differential regulation of the *Chlamydomonas* *Nar1* gene family by carbon and nitrogen. *Protist* **157**: 421–433.
- Maruyama-Nakashita, A., Inoue, E., Watanabe-Takahashi, A., Yamaya, T., and Takahashi, H. (2003). Transcriptome profiling of sulfur-responsive genes in *Arabidopsis* reveals global effects of sulfur nutrition on multiple metabolic pathways. *Plant Physiol.* **132**: 597–605.
- Maruyama-Nakashita, A., Nakamura, Y., Tohge, T., Saito, K., and Takahashi, H. (2006). *Arabidopsis* SLIM1 is a central transcriptional regulator of plant sulfur response and metabolism. *Plant Cell* **18**: 3235–3251.
- Mazel, D., and Marliere, P. (1989). Adaptive eradication of methionine and cysteine from cyanobacterial light-harvesting proteins. *Nature* **341**: 245–248.
- Merchant, S.S., et al. (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**: 245–250.
- Michelet, L., Zaffagnini, M., Marchand, C., Collin, V., Decottignies, P., Tsan, P., Lancelin, J.M., Trost, P., Miginiac-Maslow, M., Noctor, G., and Lemaire, S.D. (2005). Glutathionylation of chloroplast thioredoxin f is a redox signaling mechanism in plants. *Proc. Natl. Acad. Sci. USA* **102**: 16478–16483.
- Michelet, L., Zaffagnini, M., Vanacker, H., Le Marechal, P., Marchand, C., Schroda, M., Lemaire, S.D., and Decottignies, P. (2008). In vivo targets of S-thiolation in *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **283**: 21571–21578.
- Moharikar, S., D'Souza, J., Kulkarni, A.B., and Rao, B.J. (2006). UV-C induced apoptotic-like cell death process in the unicellular chlorophyte *Chlamydomonas reinhardtii*. *J. Phycol.* **42**: 423–433.
- Moharikar, S., D'Souza, J.S., and Rao, B.J. (2007). A homologue of the defender against the apoptotic death gene (*dad1*) in UV-exposed *Chlamydomonas* cells is downregulated with the onset of programmed cell death. *J. Biosci.* **32**: 261–270.
- Moroney, J.V., and Ynalvez, R.A. (2007). Proposed carbon dioxide

- concentrating mechanism in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **6**: 1251–1259.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B.** (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**: 621–628.
- Moseley, J., and Grossman, A.R.** (2009). Phosphorus limitation from the physiological to the genomic. In *The Chlamydomonas Sourcebook*, E.H. Harris, D. Stern, and G.B. Witman, eds (Amsterdam: Elsevier), pp. 189–216.
- Moseley, J.L., González-Ballester, D., Pootakham, W., Bailey, S., and Grossman, A.R.** (2009). Genetic interactions between regulators of *Chlamydomonas* phosphorus and sulfur deprivation responses. *Genetics* **181**: 889–905.
- Mueller, E.G.** (2006). Trafficking in persulfides: Delivering sulfur in biosynthetic pathways. *Nat. Chem. Biol.* **2**: 185–194.
- Mus, F., Dubini, A., Seibert, M., Posewitz, M.C., and Grossman, A.R.** (2007). Anaerobic acclimation in *Chlamydomonas reinhardtii*: Anoxic gene expression, hydrogenase induction, and metabolic pathways. *J. Biol. Chem.* **282**: 25475–25486.
- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M.** (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**: 1344–1349.
- Navakoudis, E., Vrentzou, K., and Kotzabasis, K.** (2007). A polyamine- and LHCII protease activity-based mechanism regulates the plasticity and adaptation status of the photosynthetic apparatus. *Biochim. Biophys. Acta* **1767**: 261–271.
- Nguyen, A.V., Thomas-Hall, S.R., Malnoe, A., Timmins, M., Mussgnug, J.H., Rupprecht, J., Kruse, O., Hankamer, B., and Schenk, P.M.** (2008). The transcriptome of photo-biological hydrogen production induced by sulphur deprivation in the green alga *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **7**: 1965–1979.
- Nikiforova, V., Freitag, J., Kempa, S., Adamik, M., Hesse, H., and Hoefgen, R.** (2003). Transcriptome analysis of sulfur depletion in *Arabidopsis thaliana*: Interlacing of biosynthetic pathways provides response specificity. *Plant J.* **33**: 633–650.
- Nikiforova, V.J., Daub, C.O., Hesse, H., Willmitzer, L., and Hoefgen, R.** (2005a). Integrative gene-metabolite network with implemented causality decipherers informational fluxes of sulphur stress response. *J. Exp. Bot.* **56**: 1887–1896.
- Nikiforova, V.J., Kopka, J., Tolstikov, V., Fiehn, O., Hopkins, L., Hawkesford, M.J., Hesse, H., and Hoefgen, R.** (2005b). Systems rebalancing of metabolism in response to sulfur deprivation, as revealed by metabolome analysis of *Arabidopsis* plants. *Plant Physiol.* **138**: 304–318.
- Ogasawara, Y., Lacourciere, G.M., Ishii, K., and Stadtman, T.C.** (2005). Characterization of potential selenium-binding proteins in the selenophosphate synthetase system. *Proc. Natl. Acad. Sci. USA* **102**: 1012–1016.
- Okazaki, Y., Shimojima, M., Sawada, Y., Toyooka, K., Narisawa, T., Mochida, K., Tanaka, H., Matsuda, F., Hirai, A., Hirai, M.Y., Ohta, H., and Saito, K.** (2009). A chloroplastic UDP-glucose pyrophosphorylase from *Arabidopsis* is the committed enzyme for the first step of sulfolipid biosynthesis. *Plant Cell* **21**: 892–909.
- Ortiz-Bermudez, P., Hirth, K.C., Srebotnik, E., and Hammel, K.E.** (2007). Chlorination of lignin by ubiquitous fungi has a likely role in global organochlorine production. *Proc. Natl. Acad. Sci. USA* **104**: 3895–3900.
- Peers, G., Truong, T.B., Ostendorf, E., Elrad, D., Grossman, A.R., Hippler, M., and Niyogi, K.K.** (2009). An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Science* **462**: 518–521.
- Pereira, Y., Lagniel, G., Godat, E., Baudouin-Cornu, P., Junot, C., and Labarre, J.** (2008). Chromate causes sulfur starvation in yeast. *Toxicol. Sci.* **106**: 400–412.
- Perez-Alegre, M., and Franco, A.R.** (1998). Resistance to L-methionine-S-sulfoximine in *Chlamydomonas reinhardtii* is due to an alteration in a general amino acid transport system. *Planta* **207**: 20–26.
- Pollock, S.V., Colombo, S.L., Prout, D.L., Jr., Godfrey, A.C., and Moroney, J.V.** (2003). Rubisco activase is required for optimal photosynthesis in the green alga *Chlamydomonas reinhardtii* in a low-CO<sub>2</sub> atmosphere. *Plant Physiol.* **133**: 1854–1861.
- Pollock, S.V., Pootakham, W., Shibagaki, N., Moseley, J.L., and Grossman, A.R.** (2005). Insights into the acclimation of *Chlamydomonas reinhardtii* to sulfur deprivation. *Photosynth. Res.* **86**: 475–489.
- Pootakham, W., González-Ballester, D., and Grossman, A.R.** (2010). Identification and regulation of plasma membrane sulfate transporters in *Chlamydomonas reinhardtii*. *Plant Physiol.* <http://dx.doi.org/10.1104/pp.110.157875>.
- Ravina, C.G., Barroso, C., Vega, J.M., and Gotor, C.** (1999). Cysteine biosynthesis in *Chlamydomonas reinhardtii*. Molecular cloning and regulation of O-acetylserine(thiol)lyase. *Eur. J. Biochem.* **264**: 848–853.
- Ravina, C.G., Chang, C.I., Tsakraklides, G.P., McDermott, J.P., Vega, J.M., Leustek, T., Gotor, C., and Davies, J.P.** (2002). The *sac* mutants of *Chlamydomonas reinhardtii* reveal transcriptional and posttranscriptional control of cysteine biosynthesis. *Plant Physiol.* **130**: 2076–2084.
- Richard, C., Ouellet, H., and Guertin, M.** (2000). Characterization of the LI818 polypeptide from the green unicellular alga *Chlamydomonas reinhardtii*. *Plant Mol. Biol.* **42**: 303–316.
- Riemenschneider, A., Wegele, R., Schmidt, A., and Papenbrock, J.** (2005). Isolation and characterization of a D-cysteine desulphydrase protein from *Arabidopsis thaliana*. *FEBS J.* **272**: 1291–1304.
- Rouhier, N., Lemaire, S.D., and Jacquot, J.P.** (2008). The role of glutathione in photosynthetic organisms: Emerging functions for glutaredoxins and glutathionylation. *Annu. Rev. Plant Biol.* **59**: 143–166.
- Sarkar, N., Lemaire, S., Wu-Scharf, D., Issakidis-Bourguet, E., and Cerutti, H.** (2005). Functional specialization of *Chlamydomonas reinhardtii* cytosolic thioredoxin h1 in the response to alkylation-induced DNA damage. *Eukaryot. Cell* **4**: 262–273.
- Sato, N., Tsuzuki, M., Matsuda, Y., Ehara, T., Osafune, T., and Kawaguchi, A.** (1995). Isolation and characterization of mutants affected in lipid metabolism of *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **230**: 987–993.
- Schloss, J.A., Silflow, C.D., and Rosenbaum, J.L.** (1984). mRNA abundance changes during flagellar regeneration in *Chlamydomonas reinhardtii*. *Mol. Cell. Biol.* **4**: 424–434.
- Scott, J., Rébeillé, F., and Fletcher, J.J.** (2000). Folic acid and folates: The feasibility for nutritional enhancement in plant foods. *J. Sci. Food Agric.* **80**: 795–824.
- Sfichi-Duke, L., Ioannidis, N.E., and Kotzabasis, K.** (2008). Fast and reversible response of thylakoid-associated polyamines during and after UV-B stress: A comparative study of the wild type and a mutant lacking chlorophyll b of unicellular green alga *Scenedesmus obliquus*. *Planta* **228**: 341–353.
- Slesak, I., Libik, M., Karpinska, B., Karpinski, S., and Miszalski, Z.** (2007). The role of hydrogen peroxide in regulation of plant metabolism and cellular signalling in response to environmental stresses. *Acta Biochim. Pol.* **54**: 39–50.
- Slocum, R.D.** (2005). Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol. Biochem.* **43**: 729–745.
- Sugimoto, K., Sato, N., and Tsuzuki, M.** (2007). Utilization of a chloroplast membrane sulfolipid as a major internal sulfur source for

- protein synthesis in the early phase of sulfur starvation in *Chlamydomonas reinhardtii*. *FEBS Lett.* **581**: 4519–4522.
- Sugimoto, K., Tsuzuki, M., and Sato, N.** (2009). Regulation of synthesis and degradation of a sulfolipid under sulfur-starved conditions and its physiological significance in *Chlamydomonas reinhardtii*. *New Phytol.* **185**: 676–686.
- Takahashi, H., Braby, C.E., and Grossman, A.R.** (2001). Sulfur economy and cell wall biosynthesis during sulfur limitation of *Chlamydomonas reinhardtii*. *Plant Physiol.* **127**: 665–673.
- Torkelson, J.D., Lynnes, J.A., and Weger, H.G.** (1995). Extracellular peroxidase-mediated oxygen consumption in *Chlamydomonas reinhardtii* (Chlorophyta). *J. Phycol.* **31**: 562–567.
- Triantaphylides, C., and Havaux, M.** (2009). Singlet oxygen in plants: Production, detoxification and signaling. *Trends Plant Sci.* **14**: 219–228.
- Triantaphylides, C., Krischke, M., Hoerberichts, F.A., Ksas, B., Gresser, G., Havaux, M., Van Breusegem, F., and Mueller, M.J.** (2008). Singlet oxygen is the major reactive oxygen species involved in photooxidative damage to plants. *Plant Physiol.* **148**: 960–968.
- Voigt, J., Deinert, B., and Bohley, P.** (2000). Subcellular localization and light-dark control of ornithine decarboxylase in the unicellular green alga *Chlamydomonas reinhardtii*. *Physiol. Plant.* **108**: 353–360.
- Vreeland, V., Waite, J.H., and Epstein, L.** (1998). Polyphenols and oxidases in substrate adhesion by marine algae and mussels. *J. Phycol.* **34**: 1–8.
- Wang, Z., Gerstein, M., and Snyder, M.** (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**: 57–63.
- Williams, B., and Dickman, M.** (2008). Plant programmed cell death: Can't live with it; can't live without it. *Mol. Plant Pathol.* **9**: 531–544.
- Wykoff, D., Davies, J., and Grossman, A.** (1998). The regulation of photosynthetic electron transport during nutrient deprivation in *Chlamydomonas reinhardtii*. *Plant Physiol.* **117**: 129–139.
- Yildiz, F., Davies, J.P., and Grossman, A.R.** (1994). Characterization of sulfate transport in *Chlamydomonas reinhardtii* during sulfur-limited and sulfur-sufficient growth. *Plant Physiol.* **104**: 981–987.
- Yoda, H., Yamaguchi, Y., and Sano, H.** (2003). Induction of hypersensitive cell death by hydrogen peroxide produced through polyamine degradation in tobacco plants. *Plant Physiol.* **132**: 1973–1981.
- Zhang, L., Happe, T., and Melis, A.** (2002). Biochemical and morphological characterization of sulfur-deprived and H<sub>2</sub>-producing *Chlamydomonas reinhardtii* (green alga). *Planta* **214**: 552–561.
- Zhang, L., and Melis, A.** (2002). Probing green algal hydrogen production. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**: 1499–1507.
- Zhang, Z., Shrager, J., Jain, M., Chang, C.W., Vallon, O., and Grossman, A.R.** (2004). Insights into the survival of *Chlamydomonas reinhardtii* during sulfur starvation based on microarray analysis of gene expression. *Eukaryot. Cell* **3**: 1331–1348.

# Correction

**David González-Ballester, David Casero, Shawn Cokus, Matteo Pellegrini, Sabeeha S. Merchant, and Arthur R. Grossman.** (2010). RNA-Seq Analysis of Sulfur-Deprived *Chlamydomonas* Cells Reveals Aspects of Acclimation Critical for Cell Survival. *Plant Cell* **22**: 2058–2084.

Table 3 has been corrected to read:

**Table 3.** Expression Levels of Selected S-Regulated Genes

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type –S/ Wild Type +S		<i>snrk2.1</i> /Wild Type (+S)		<i>snrk2.1</i> / Wild Type (–S)		Wild Type	Wild Type	<i>snrk2.1</i> –S	<i>snrk2.1</i> +S
		S	Q	S	Q	S	Q	–S	+S	–S	+S
Sulfur acquisition and assimilation											
205496	<i>ARS1</i> , arylsulfatase (ext.)	<b>10.5</b>	<b>9.5</b>	–1.2	–0.5	<b>–16.1</b>	<b>–12.0</b>	<b>2717</b>	2	0	1
55757	<i>ARS2</i> , arylsulfatase (ext.)	<b>8.2</b>	<b>12.2</b>	1.0	<b>–2.9</b>	<b>–11.1</b>	<b>–15.1</b>	84	0	0	1
205502	<i>SLT1</i> , Na <sup>+</sup> / SO <sub>4</sub> <sup>2–</sup> transporter type (p)	<b>6.9</b>	<b>7.9</b>	<b>–6.3</b>	<b>–7.3</b>	<b>–15.0</b>	<b>–20.4</b>	<b>2341</b>	20	0	0
205501	<i>SLT2</i> , Na <sup>+</sup> / SO <sub>4</sub> <sup>2–</sup> transporter type (p)	<b>6.3</b>	<b>9.3</b>	<b>–2.5</b>	–1.0	<b>–9.8</b>	<b>–10.3</b>	<b>1339</b>	17	1	3
150514	<i>SULTR2</i> , H <sup>+</sup> /SO <sub>4</sub> <sup>2–</sup> transporter type (p)	<b>5.3</b>	<b>7.7</b>	<b>–4.5</b>	<b>–3.9</b>	<b>–6.4</b>	<b>–7.5</b>	<b>152</b>	4	2	0
196910	<i>ATS1</i> , ATP sulfurylase	<b>1.6</b>	<b>1.6</b>	–1.2	–1.3	<b>–3.0</b>	<b>–2.9</b>	<b>2137</b>	<b>723</b>	<b>275</b>	<b>305</b>
133924	<i>ATS2</i> , ATP sulfurylase (chl)	<b>2.8</b>	<b>3.6</b>	0.0	0.0	<b>–4.4</b>	<b>–3.6</b>	<b>254</b>	36	12	37
128906	<i>SULP1</i> , chl. SO <sub>4</sub> <sup>2–</sup> transport system	<b>2.4</b>		–0.6		<b>–3.4</b>		<b>147</b>	28	14	18
116547	<i>SULP2</i> , chl. SO <sub>4</sub> <sup>2–</sup> transport system	1.5		–0.7		<b>–2.2</b>		56	20	12	12
182359	<i>SBP1</i> , chl. SO <sub>4</sub> <sup>2–</sup> transport system	1.5		–1.1		<b>–2.7</b>		79	28	12	13
194724	<i>SABC</i> (CysA), chl. SO <sub>4</sub> <sup>2–</sup> transport system	<b>2.6</b>		–0.5		<b>–2.7</b>		<b>144</b>	24	23	17
184419	<i>APK</i> , APS kinase (chl.)	0.4	1.3	0.5	–0.7	<b>–1.7</b>	<b>–1.6</b>	39	29	12	41
131444	<i>APR</i> (MET16), APS sulforeductase (chl.)	0.5	1.4	–0.9	0.8	<b>–2.3</b>	<b>–0.6</b>	<b>285</b>	<b>197</b>	57	<b>105</b>
206154	<i>SIR1</i> , ferredoxin-sulfite reductase (chl.)	<b>2.0</b>		–1.1		<b>–2.3</b>		<b>124</b>	31	26	14
205485	<i>SIR2</i> , ferredoxin-sulfite reductase	<b>2.2</b>		–0.7		<b>–2.2</b>		<b>597</b>	<b>126</b>	<b>132</b>	80
189320	<i>OASTL4a</i> , O-acetylserine(thio)lyase (chl.)	<b>2.9</b>	<b>5.6</b>	–0.3	–0.3	<b>–5.0</b>	<b>–6.6</b>	<b>976</b>	<b>128</b>	31	<b>103</b>
205985	<i>SAT1a</i> , serineacetyl transferase (chl.)	<b>5.1</b>	<b>4.2</b>	–0.7	0.1	<b>–4.3</b>	<b>–4.3</b>	<b>794</b>	23	40	14
24268	<i>CGS1</i> (METB), cystathionine g-synthase (chl.)	<b>–1.6</b>		–0.6		1.4		23	71	63	49
127384	<i>THS1</i> , Thr synthase (chl.)	–1.4		–0.7		–0.3		54	<b>138</b>	42	84
196483	<i>METC</i> , cystathionine b-lyase (chl.)	–0.5		–0.5		1.3		24	34	57	24
154307	<i>METE</i> , Met synthase	<b>–3.7</b>		–0.6		0.2		46	<b>606</b>	54	<b>395</b>
76715	<i>METH1</i> , Met synthase (cobalamin)	<b>–2.0</b>		–0.9		–0.2		33	<b>134</b>	29	72
195332	<i>METH2</i> , Met synthase (cobalamin)	<b>–2.8</b>		–0.8		1.1		30	<b>211</b>	65	<b>123</b>

(Continued)

**Table 3** (continued).

ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type -S/ Wild Type +S		snrk2.1/Wild Type (+S)		snrk2.1/ Wild Type (-S)		Wild Type -S	Wild Type +S	snrk2.1 -S	snrk2.1 +S
		S	Q	S	Q	S	Q				
182408	<i>METM</i> , S-adenosylmethionine synthetase	-2.1		-0.2		-0.3		526	2288	420	1976
129593	<i>SAH1</i> , S-adenosylhomocysteine hydrolase	-3.1		-0.4		-0.9		152	1298	83	952
183928	<i>DCA1</i> , S-adenosylmethionine decarboxylase	-1.8		-0.3		1.5		54	184	149	146
206104	<i>AOT2</i> , amino acid/polyamine transporter (p)	4.1	3.6	-5.5	-2.4	-9.8	-6.3	36	2	0	0
206105	<i>AOT4</i> , amino acid/polyamine transporter (p)	3.9	4.0	-8.0	-8.8	-13.0	-13.1	311	21	0	0
185190	<i>THI4a</i> , thiazole biosynthetic enzyme	-3.2		-0.1		-2.8		203	1859	29	1682
196899	<i>THI4b</i> , THI4 regulatory protein	-3.2		-0.1		-2.8		188	1715	27	1557
192720	<i>THICa</i> , hydroxymethylpyrimidine phosphate synthase	-3.9		-0.4		-3.3		6	95	1	74
196900	<i>THICb</i> , hydroxymethylpyrimidine phosphate synthase	-3.9		-0.4		-3.4		6	92	1	72
142289	<i>BIOB1</i> , biotin synthase	-2.5		1.1		-0.5		2	10	1	23
97943	<i>BIOB2</i> , biotin synthase	-1.9		-0.1		-1.6		23	86	8	77
181975	<i>GSH1</i> , γ-glutamylcysteine synthetase (chl.)	0.1		0.0		0.1		35	33	38	34
189020	<i>GSH2</i> , glutathione synthetase (chl.)	-0.3		-0.3		-1.4		20	25	8	20
170508	<i>GTP1</i> , γ-glutamyl transpeptidase	0.2		0.7		0.7		25	23	41	37
Redistribution and recycling of S											
143892	<i>CDO1</i> , Cys dioxygenase	4.7	4.5	-1.4	-4.0	-7.3	-8.4	138	5	1	2
127464	<i>TAUD1</i> , taurin dioxygenase	7.1	5.6	-3.5	-5.3	-11.7	-12.5	1417	10	0	1
77600	<i>TAUD2</i> , taurin dioxygenase	5.2	6.6	-0.4	0.9	-5.1	-6.4	632	17	19	13
59800	<i>SUOX1</i> , sulfite oxidase, mit.	2.0		-0.2		-1.8		149	36	42	31
183511	<i>RDP3</i> , small rodhanase	4.1	4.4	-0.8	-2.7	-4.0	-6.6	369	21	24	12
132010	<i>TST</i> , mercaptopyruvate sulfurtransferase	0.2		0.4		-1.5		22	19	8	26
167884	<i>CSD2 (NIFS2)</i> , Cys desulfurase	1.6		-0.2		-2.0		43	14	10	12
98369	D-Cys desulfhydrase Lipid metabolism	2.2		0.2		-2.5		37	8	7	9
144554	Esterase/lipase/thioesterase	3.1		-0.5		-1.6		65	7	21	5
27658	<i>SQD1</i> , UDP-sulfoquinovose synthase	1.7		-0.9		-4.1		336	102	19	57
206163	<i>SQD2</i> , sulfolipid synthase	1.5		-0.3		-5.1		53	19	2	16
196477	<i>LPB1</i> , low phosphate bleaching	4.1	4.2	-1.6	-0.7	-5.5	-5.8	1711	98	39	33
Photosynthesis											
Stress-related genes											
148916	<i>ELIP3</i> , chlorophyll <i>a/b</i> binding protein	2.9		0.0		0.0		42	6	42	6
184724	<i>LHCSR1</i> , light-harvesting stress-related protein (chl.)	1.8	2.2	0.6	1.2	-1.6	-2.1	596	174	202	257
184731	<i>LHCSR2</i> , light-harvesting stress-related protein (chl.)	2.1	4.6	1.4	5.9	0.4	2.0	244	56	315	148
184730	<i>LHCSR3</i> , light-harvesting stress-related protein (chl.)	2.3	5.5	1.1	7.0	-0.1	3.0	611	128	564	275
Major light-harvesting proteins of PSII											
185533	<i>LHCBM1</i> (chl.)	-0.6	-2.9	0.2	-0.3	-1.1	-7.2	5425	8019	2584	9327
184067	<i>LHCBM2</i> (chl.)	-0.2	-2.9	0.2	-0.3	-0.6	-1.9	2705	3188	1740	3774
195162	<i>LHCBM3</i> (chl.)	-1.7	-4.5	0.5	0.1	-0.2	-5.5	676	2262	606	3112

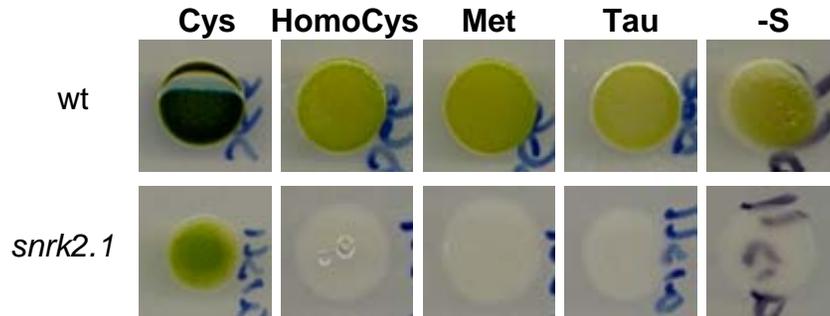
(Continued)

**Table 3** (continued).

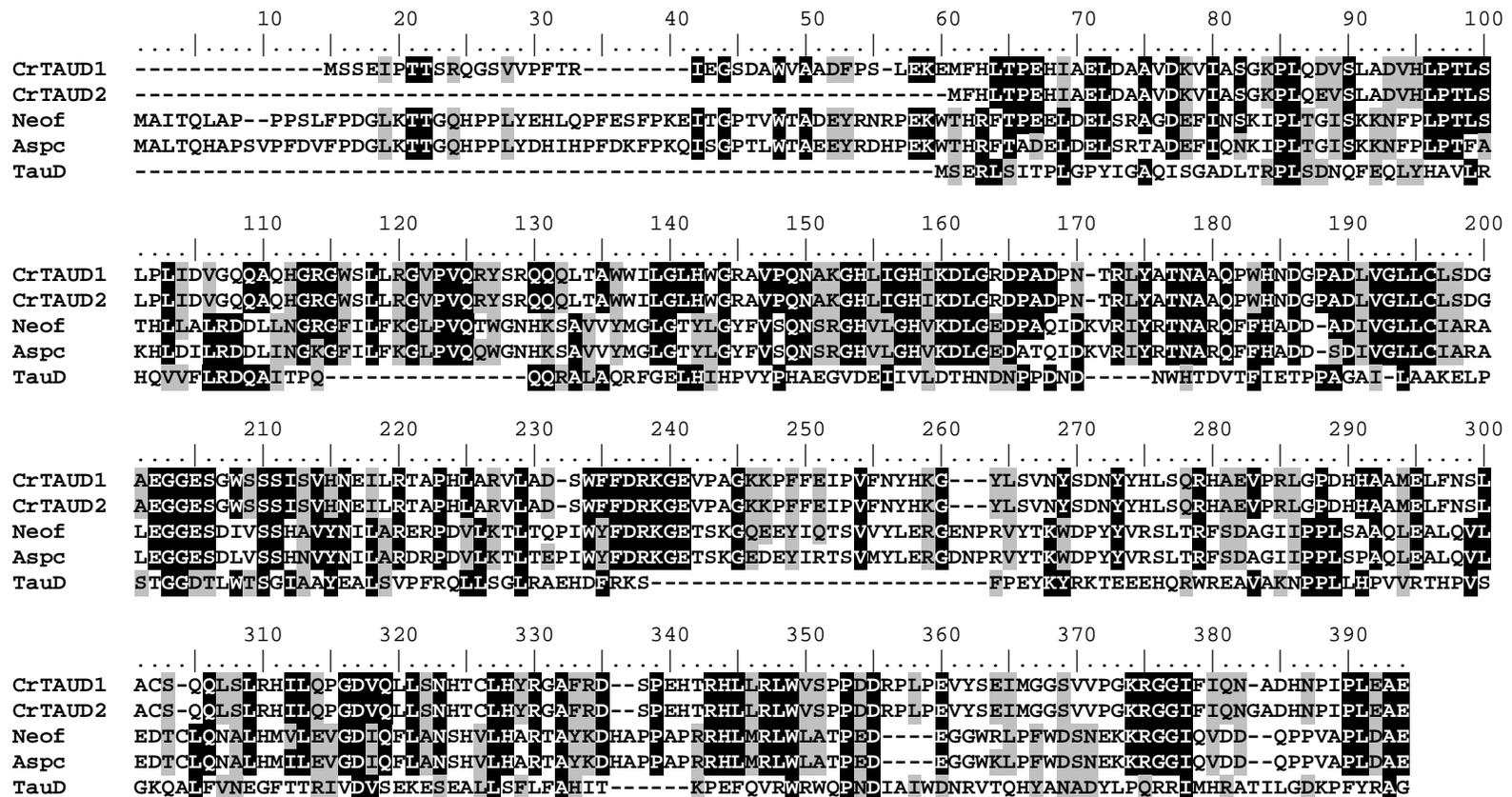
ID	Name/Description	Fold Change Expression (Log <sub>2</sub> )						Transcript Abundance (RPKM) <sup>a</sup>			
		Wild Type -S/ Wild Type +S		snrk2.1/Wild Type (+S)		snrk2.1/ Wild Type (-S)		Wild	Wild	snrk2.1	snrk2.1
		S	Q	S	Q	S	Q	Type -S	Type +S	-S	+S
191690	<i>LHCBM4</i> (chl.)	-0.3	<b>-3.5</b>	0.3	0.2	<b>-2.2</b>	-1.2	<b>747</b>	<b>951</b>	<b>163</b>	<b>1160</b>
184775	<i>LHCBM5</i> (chl.)	-1.4	<b>-3.4</b>	0.6	0.4	-2.6	<b>-6.8</b>	<b>564</b>	<b>1485</b>	95	<b>2292</b>
184490	<i>LHCBM6</i> (chl.)	-1.1	<b>-2.6</b>	0.0	0.1	0.0	-1.3	<b>1177</b>	<b>2448</b>	<b>1163</b>	<b>2494</b>
184071	<i>LHCBM7</i> (chl.)	-0.9	<b>-4.3</b>	0.4	0.1	<b>-2.3</b>	<b>-5.6</b>	<b>3673</b>	<b>7052</b>	<b>732</b>	<b>9094</b>
205752	<i>LHCBM8</i> (chl.)	-0.5	<b>-7.5</b>	0.2	-0.6	<b>-3.0</b>	<b>-1.6</b>	80	<b>114</b>	10	<b>129</b>
184479	<i>LHCBM9</i> (chl.)	<b>10.0</b>	<b>8.8</b>	<b>-2.6</b>	<b>-10.1</b>	<b>-12.8</b>	<b>-20.5</b>	<b>21455</b>	21	3	3
Carbon metabolism											
196354	<i>SHMT2</i> , Ser hydroxymethyltransferase 2	<b>-3.0</b>		-0.4		0.5		31	<b>254</b>	43	<b>196</b>
101528	6-phosphogluconolactonase-like protein	<b>2.0</b>		-0.2		-1.3		<b>401</b>	<b>100</b>	<b>160</b>	85
163301	<i>GLD2-C</i> , Glc-6-phosphate dehydrogenase	<b>2.3</b>		-0.3		<b>-2.1</b>		77	16	18	13
192597	<i>GND1a</i> , 6-phosphogluconate dehydrogenase	<b>1.9</b>		-0.2		<b>-1.6</b>		<b>308</b>	83	<b>101</b>	74
158911	(decarboxylating) (chl.)										
158911	<i>GND1b</i> , 6-phosphogluconate dehydrogenase	<b>1.9</b>		-0.2		<b>-1.6</b>		<b>313</b>	86	<b>102</b>	74
163238	(decarboxylating) (cyt.)										
163238	<i>RIBA</i> , bifunctional GTP cyclohydrolase II/3,4-dihydroxy-2butanone-4-phosphate synthase	<b>1.7</b>		-0.7		<b>-2.4</b>		83	26	15	16
196834	<i>CAH9</i> , carbonic anhydrase (cyt.?)	<b>5.0</b>		-1.1		<b>-6.8</b>		<b>1952</b>	62	18	28
Folate metabolism											
206121	<i>MTDH/MTCH2</i> , methylenetetrahydrofolate dehydrogenase/ methylenetetrahydrofolate cyclohydrolase	<b>-1.6</b>		-0.2		<b>-1.5</b>		13	40	5	35
127560	<i>FTHFS</i> , 10-formyltetrahydrofolate synthetase	<b>-2.5</b>		0.3		<b>-1.8</b>		10	59	3	71
30522	<i>ADCL1</i> , aminotransferase related to 4-amino-4-deoxychorismate lyase	<b>-1.7</b>		-0.2		0.2		23	73	26	65
111330	<i>MTHFR</i> , 5,10-methylenetetrahydrofolate reductase	<b>-3.1</b>		-0.1		0.1		11	92	12	86
182461	<i>GGH1</i> , $\gamma$ -glutamyl hydrolase	<b>3.9</b>		-0.6		<b>-6.2</b>		36	2	0	2
Extracellular/cell wall proteins											
168785	<i>HAP1</i> , vanadium haloperoxidase	<b>7.9</b>	<b>7.8</b>	-0.2	<b>-2.0</b>	<b>-8.4</b>	<b>-11.7</b>	<b>739</b>	3	2	3
143696	<i>HAP2</i> , vanadium haloperoxidase	<b>8.4</b>	<b>7.1</b>	-0.1	-0.6	<b>-11.6</b>	<b>-11.5</b>	<b>1044</b>	3	0	3
182794	<i>HAP3</i> , vanadium haloperoxidase	<b>7.1</b>	<b>6.5</b>	<b>-3.2</b>	-1.0	<b>-11.1</b>	<b>-7.2</b>	<b>1456</b>	10	1	1
130684	<i>ECP76</i> , extracellular protein	<b>10.7</b>	<b>10.3</b>	-0.8	-0.4	<b>-13.1</b>	<b>-11.8</b>	<b>899</b>	1	0	0
137329	<i>ECP88</i> , extracellular protein	<b>13.4</b>		<b>-2.3</b>		<b>-15.9</b>		<b>2295</b>	0	0	0
119420	<i>ECP61</i> , extracellular protein	<b>8.5</b>		-0.6		<b>-11.0</b>		<b>268</b>	1	0	0
194201	<i>ECP56</i> , extracellular protein	<b>8.6</b>		<b>-4.2</b>		<b>-13.0</b>		<b>317</b>	1	0	0

Fold change (log<sub>2</sub> values) in transcript levels under specified conditions as determined by RNA-seq (S) and, when available, by qRT-PCR (Q). Values above 1.5 or below -1.5 are marked in bold type. Some genes for which expression is not affected by -S conditions are included for informative purposes. Putative subcellular localizations are denoted in parentheses after the transcript names as: cyt., cytosolic; mit., mitochondria; chl., chloroplast; and ext., extracellular. ID, JGI protein accession number; p, putative.

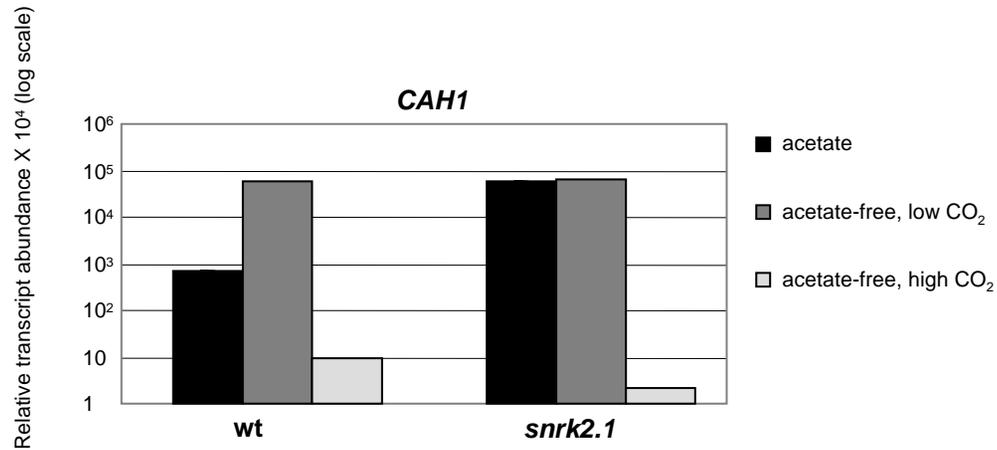
<sup>a</sup>Transcript abundance obtained from RNA-seq is indicated as RPKMs (see Methods).



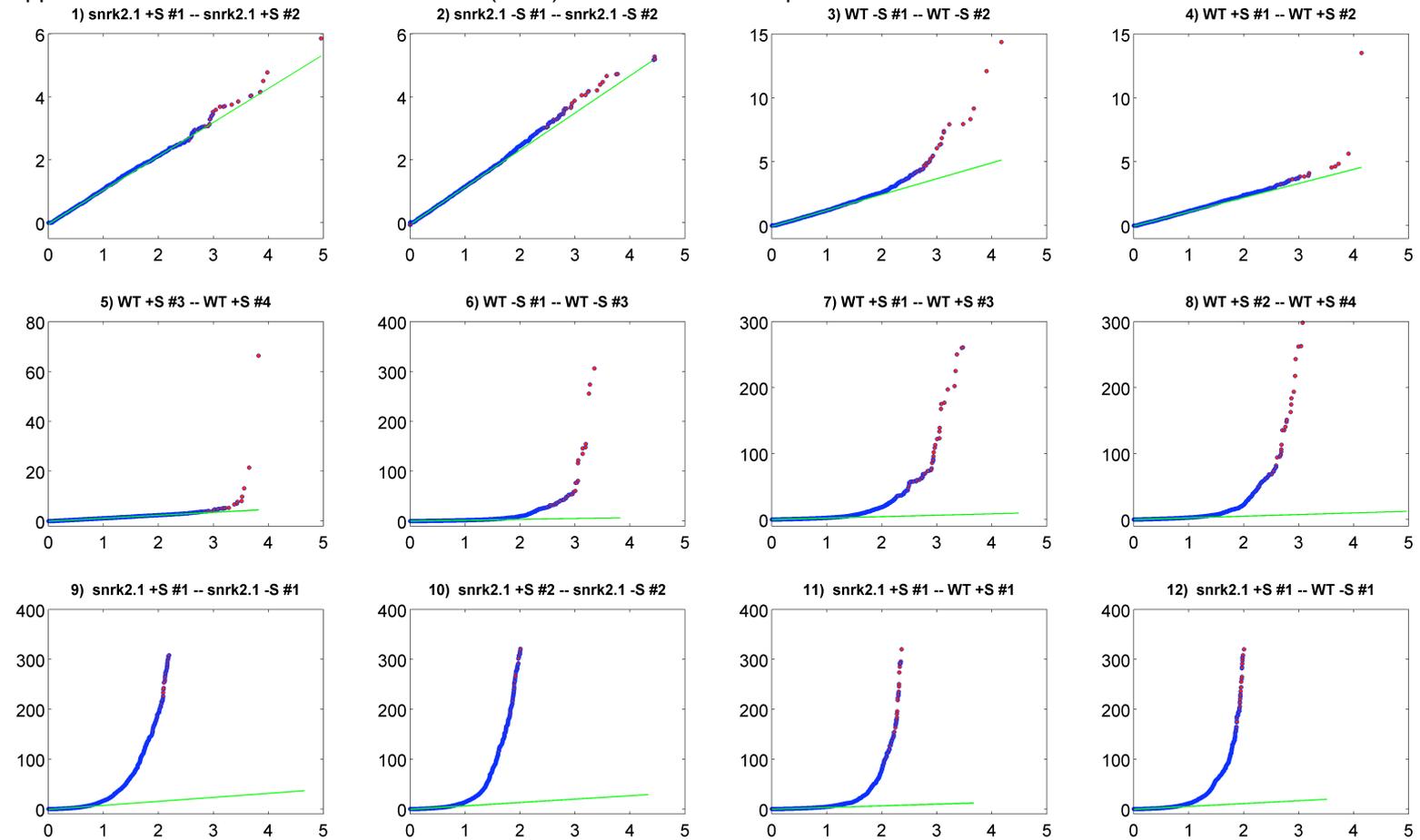
**Supplemental Figure 1. Chlamydomonas growth with different S-containing amino acid as sole S source.** Equal number of cells from the wt and *snrk2.1* strains were inoculated on TAP-S agarose (0.8 %) plates supplemented with 100 mg/L of the different S-containing amino acids. Plates were maintained under continuous light for 11 days. Cys, cysteine; Hcys, homocysteine; Met, methionine; Tau, taurine; -S, absence of any S source.



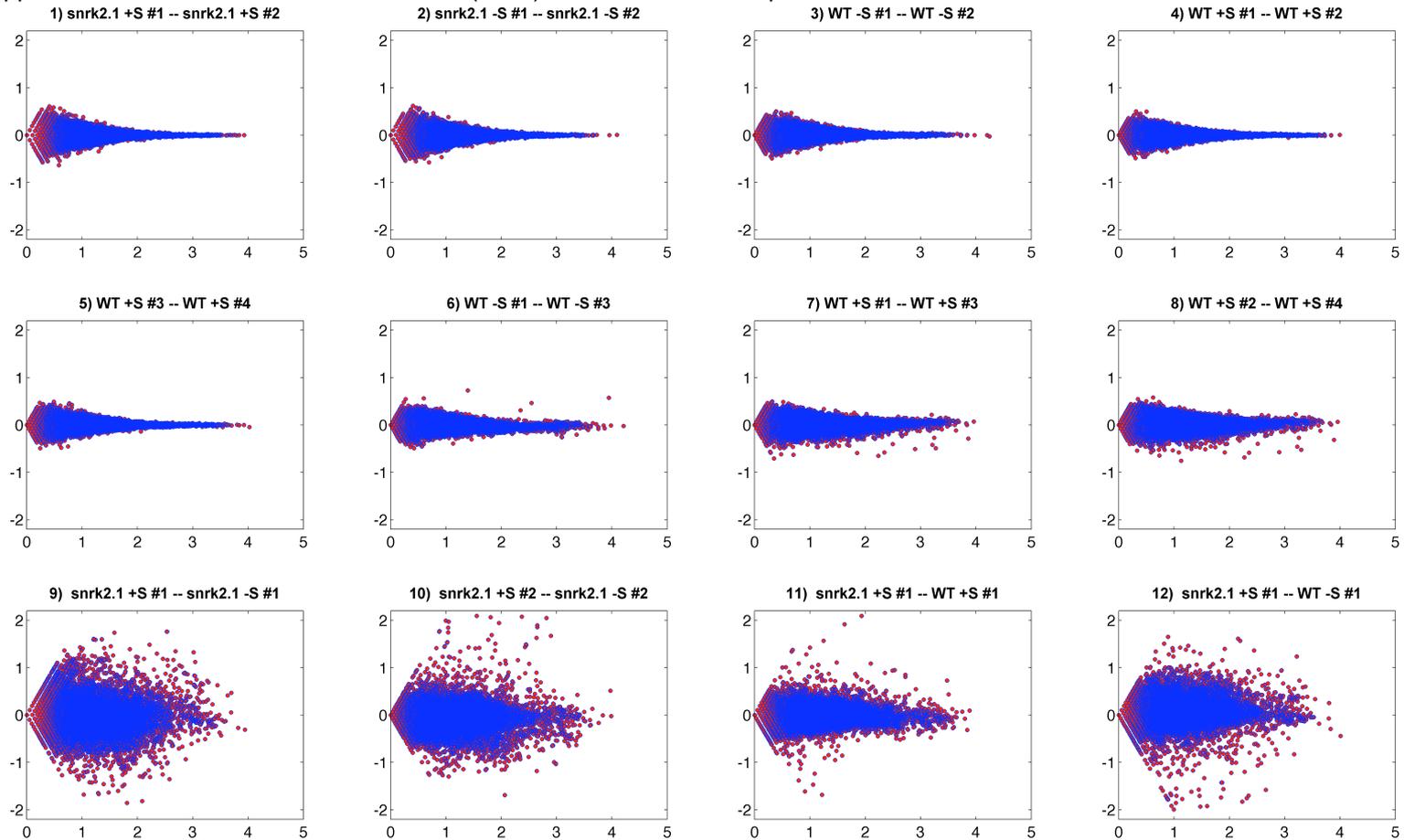
**Supplemental Figure 2. Alignment of putative taurine dioxygenases (TAUD).** Amino acid sequence alignments of *Chlamydomonas* TAUD1 and TAUD2, *Aspergillus clavatus* (Aspc) XP\_001272674.1, *Neosartorya fischeri* (Neof) XP\_001260146.1 and *E. coli* TauD. The black and grey boxes indicate identical and similar amino acids, respectively. Alignment was performed using BioEdit 7.0.5.3 software.



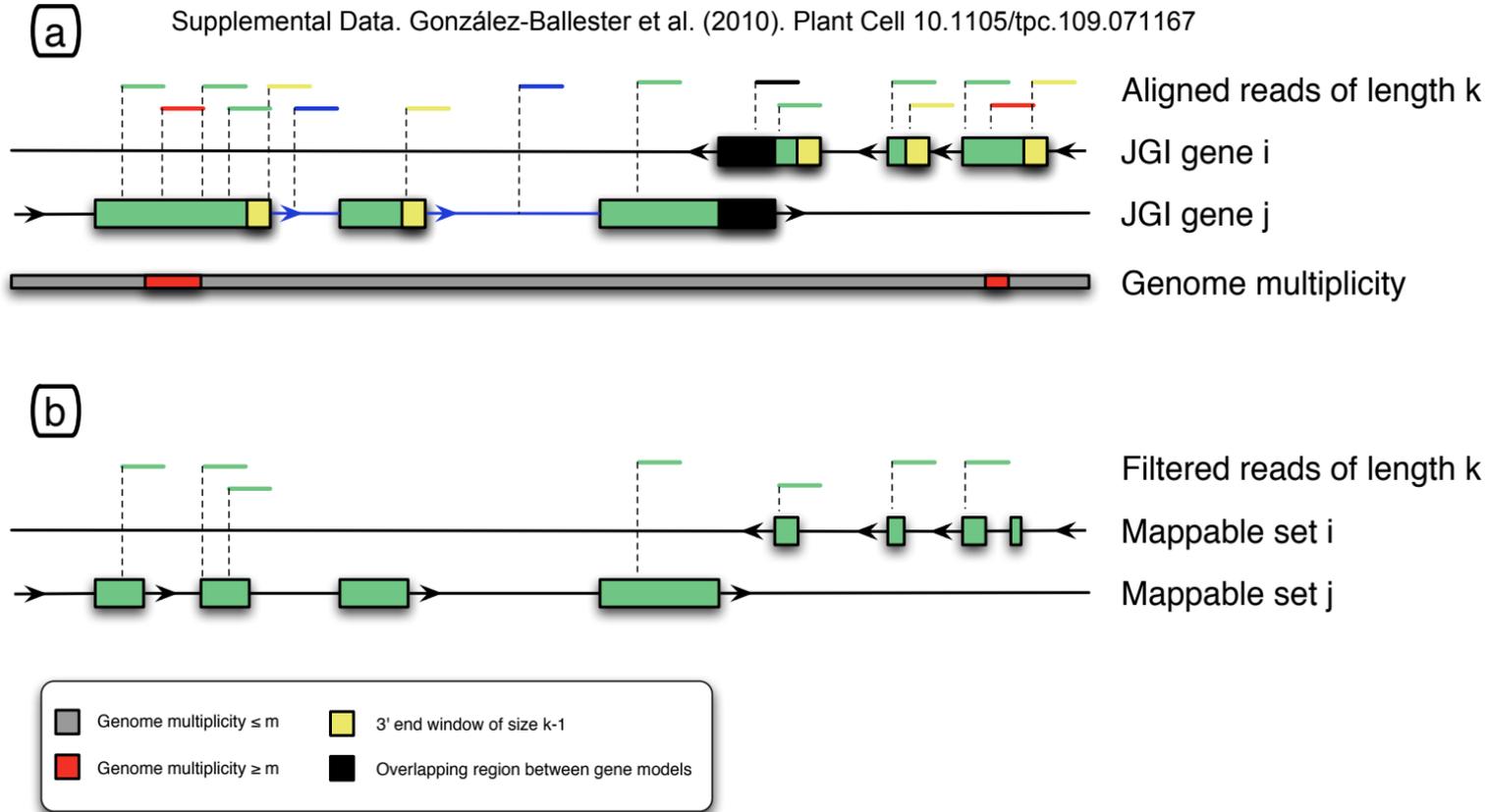
**Supplemental Figure 3. *CAH1* transcript accumulation in wt and *snrk2.1* cells under fixed and inorganic carbon conditions (based on RT-qPCR).** RNA samples were isolated after the cells were exposed to media with different inorganic carbon conditions for 6 h. Levels of individual transcripts are given as relative fold abundance with respect to the housekeeping control gene (*CBLP*) and then were multiplied by a factor of 10<sup>4</sup>. Graphics are in log scale.



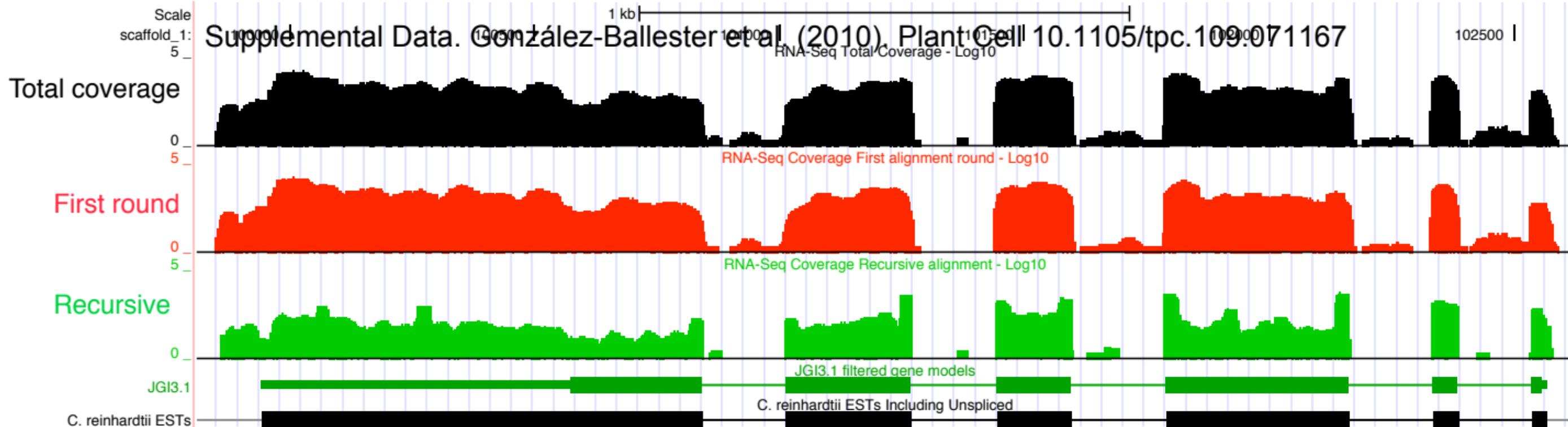
**Supplemental Figure 4. qq-plots comparing p-values from selected pairwise comparisons of the total number of hits per gene from RNA-seq individual lanes** (see Supplemental Table 3 for details). For each comparison, the negative log p-values (y axes) are compared against the quantiles from a uniform distribution in the interval [0,1] (x axis). The first two panels correspond to the same biological sample sequenced in different lanes of the same flow cell. Panels 3 to 8 compare the same biological sample when sequenced in different flow cells. Deviation from linearity (green line) indicates a flow cell effect in some cases. The last four panels are included for comparison purposes and refer to the same type of plot for lanes from different biological samples. In this case there is a much more pronounced skew of the p-value distribution.



**Supplemental Figure 5. Logarithmic scatterplots of normalized expression fold-changes (y axes) vs. overall expression measures (x axes).** The plots were generated from the same set of pair-wise comparisons included in **Suppl. Fig. 3**. Biological effects (last four panels) downplay the relevance of lane and flow cell effects.



**Supplemental Figure 6. Definition of mappable sets and assignment of  $k$ -hits per gene.** a) Aligned reads of length  $k$  (colored lines) that overlap with two different genes ( $i, j$ ) are shown along with their corresponding hits (dashed lines). Two different types of hits are filtered: those corresponding to complete (blue) or partial (yellow) alignments embedded in intergenic and intronic regions, and ambiguous hits on regions with overlapping genes (black) or with genome multiplicity higher than a predefined threshold  $m$  (red). b) The remaining coding sequence after the previous filters are applied is denoted as the *mappable set*, and its length is denoted as *gene  $m$ -mappability*.



**Supplemental Figure 7. Coverage data for a 3 kb region of the GAP3 locus** The red and green tracks were computed for unique hits from either the first or recursive alignment round, respectively. The black track shows total coverage from all the alignment rounds. The recursive alignment contributes to the total coverage mostly at intron-exon junctions, which allows for confident identification of small exons.

SINGLET OXYGEN STRESS-RELATED GENES		Fold change expression (log <sub>2</sub> )			Transcript abundance (RPKM)			
		A Ledford et al	B wt -S/+S	C <i>snrk2.1</i> -S/+S	wt -S	wt +S	<i>snrk2.1</i> -S	<i>snrk2.1</i> +S
<b>Up-regulated</b>								
193661	<i>GSTS1</i> , glutathione-S-transferase	<b>1.8</b>	0.1	<b>4.2</b>	37	34	<b>494</b>	28
143122	<i>GPXH</i> ( <i>GPX5</i> ), Glutathione peroxidase, cyt.	<b>6.8</b>	0.7	<b>2.9</b>	<b>125</b>	78	<b>502</b>	68
196808	<i>INDA</i> , indigoidine synthase (p)	<b>1.7</b>	0.9	<b>2.5</b>	11	6	27	5
195887	<i>TRXh1</i> , thioredoxin, cyt.	<b>1.6</b>	-0.2	<b>1.6</b>	<b>190</b>	<b>216</b>	<b>643</b>	<b>210</b>
196005	<i>PHC8</i> , cell wall protein pherophorin	<b>1.7</b>	0.6	<b>1.6</b>	55	36	<b>151</b>	50
190453	Unknown	<b>1.6</b>	0.9	1.2	78	41	<b>129</b>	55
<b>Down-regulated</b>								
185190	<i>THI4a</i> , thiazole biosynthetic enzyme	<b>-2.0</b>	<b>-3.2</b>	<b>-5.8</b>	<b>203</b>	<b>1859</b>	29	<b>1682</b>
58140	<i>AGS1</i> , arginosuccinate synthase	<b>-1.7</b>	-0.5	<b>-3.3</b>	73	<b>100</b>	10	97
158519	Unknown	-1.5	0.8	<b>-2.9</b>	<b>259</b>	<b>149</b>	48	<b>355</b>
115491	Acetate transporter (p)	<b>-1.5</b>	<b>-1.7</b>	<b>-1.8</b>	<b>699</b>	<b>2285</b>	<b>228</b>	<b>806</b>
183351	<i>OASTL4a</i>	<b>-1.9</b>	<b>2.9</b>	<b>-1.7</b>	<b>976</b>	<b>128</b>	31	<b>103</b>
24120	<i>CAH1</i> , carbonic anhydrase	<b>-2.9</b>	<b>-2.1</b>	<b>-1.5</b>	<b>328</b>	<b>1375</b>	<b>1727</b>	<b>5057</b>
128227	<i>CMPS1</i> , carbamoyl phosphate synthase, small s.(p)	<b>-1.6</b>	-0.2	-1.0	<b>112</b>	<b>131</b>	78	<b>153</b>
189430	<i>CCP1</i> , low-CO <sub>2</sub> -inducible chl. envelope protein	<b>-1.7</b>	-0.7	-0.7	<b>139</b>	<b>230</b>	<b>529</b>	<b>830</b>

**Supplemental Table 1. Chlamydomonas singlet oxygen stress-related genes and their expression in wt and *snrk2.1* mutant in -S conditions.** All genes found to increase or decrease during acclimation to singlet oxygen (Ledford et al., 2007) showed the same pattern of expression as in the *snrk2.1* mutant exposed to -S conditions. Original ratios of expression levels measured in Ledford et al. (2007) are shown in column A. Fold change (in log<sub>2</sub> values) in transcript levels between specified conditions, as determined by RNA-seq data, are indicated (columns B and C). Others Table details are given in **Table 3**.

<i>Gene</i>		<i>Sequence (5' - 3')</i>	<i>Gene</i>		<i>Sequence (5' - 3')</i>
<i>APK1</i>	F	TGGTGATGGGGGCTGGACATAGT	<i>LHCBM7</i>	F	GCGTGATTGAGCGGAGATGAAGAG
	R	GACCCGGCATTCCAAGTTCTCCT		R	CCGCAAGCAAAACCCGAACAC
<i>APR1</i>	F	AGGGAGCGGCTGGATGTGTTGT	<i>LHCBM8</i>	F	GGTTTGTGCTGGGGCTCGTA
	R	TCTGCCCTCTGTCGCTCATCCA		R	GCTCCGCACCACTACCTCACCA
<i>ARS1</i>	F	CGCGCCGTCACTTGTGTTGTTG	<i>LHCBM9</i>	F	CCGCCTGGCCATGTTCTCGTC
	R	GCCCACCTCTTTACCCAGCACCTC		R	CCGTGGGCCCTATCCGTGGTA
<i>ARS2</i>	F	GTACTGCCGCGTGCCTGATTCC	<i>LPB1</i>	F	TAAGACACCGTTGGCCGTTGAAAG
	R	CCTAAACATTTGGCTCCGCAGTCC		R	TAGTCGGGCCCAGGGGTAGGTAAG
<i>ATS1</i>	F	GACGACATCCCCGAGTGGTTC	<i>OAST4</i>	F	GGCCCTCCTGGAAGCTGAATCA
	R	GGCCAAGCCCGTCAACAGC		R	CCCCACCCACCCATTAGTAGTC
<i>ATS2</i>	F	TGCCTCAAGAGCGTGCCTCAG	<i>PWR1</i>	F	GATGCGCATGCTCGGCTTTC
	R	CCCGGCAGCGGGAACCATT		R	ACGCGCCCTTGCCTATCTCTA
<i>CBLP</i>	F	CTTCTCGCCCATGACCAC	<i>RDP1</i>	F	TACCATCGGGCGTGTTCAGTCC
	R	CCCACCAGGTTGTTCTTCAG		R	CTCGGGCGCGTTGTCCAGAC
<i>CAH1</i>	F	CGCGAACCCCGATGCCTACA	<i>RDP3</i>	F	GACCCCGCGCAGAAGTGATTG
	R	AAGCGCGGGTCGTCAAGAACA		R	CCGTACCATTGCGTCCCTGTCT
<i>CD01</i>	F	GCGTTGCGGGCCTATGGTAG	<i>SAT1</i>	F	TGCGGCACCCTTTCCAGAG
	R	GCGCAGCAGTGGGTGTCGTT		R	CACCCGATCACACGCAAACCTCA
<i>CD02</i>	F	CCCAAGTCTGCGTGCTTCTAGTC	<i>SBDP</i>	F2	GGACGGCAGCATCATGGTGAGC
	R	ACGCGGCTCACACCTGTCT	R2	R2	TCCACACGCCCTTGACCTTGAG
<i>ECP76</i>	F	CCTCGCTCTCCTCGCTGCTG	<i>SIR2</i>	F	CCTTCATCACCGAGGACTCGG
	R	CGGCCGACTTGGGTAATTGC		R	CCGTCCGCAGGCCTGAT
<i>HAP1</i>	F	CCAAGCCCAGGGATGAGCA	<i>SLT1</i>	F	ACGGGTCTTCGAGCGAATTGC
	R	CGATGGACGCGGATGGAAGAA		R	CGACTGCTTACGCAACAATCTTGG
<i>HAP2</i>	F	AAGGATTGTGTCAGGCTCGTCTCG	<i>SLT2</i>	F	GTACGGAGTTCCTTACGCGC
	R	TTCCACCCGACGCTAACCACA		R	TTCTTCGCCACCGATGAGC
<i>HAP3</i>	F	GCCGGCAAGACCTTCGTGGAC	<i>SULTR2</i>	F	ACGTGGCATGCAGCTCAT
	R	GCCCGGCCCTTGCTATGTGC		R	CTTGCCACTTTGCCAGGT
<i>LHCRS1</i>	F	CTGAGAGGGCCAAAGAATAAGAGC	<i>TAUD1</i>	F	CGGAAACTCCACGCCGCAAA
	R	CCATGAGAGGGGAAGATACGA		R	CGTCAGCCGCCCTTCACTCA
<i>LHCRS2</i>	F	GGAAGCAGCAAAACGCACACA	<i>TAUD2</i>	F	ACTGGGGGCGGATTTGTGTATG
	R	GGGACAGTCCCAGCTCCTATGC		R	CTGTGATCCAGGACGCCAGAGTC
<i>LHCRS3</i>	F	ACTTTTGAGTTGCTGCTGCGATA	<i>AOT2</i>	F	CACGTAATCCTGCTTGGTGGTGC
	R	CACCCCTGCTCTCGCCTGTT		R	AGCGTGGGCCGGTTCTGGTA
<i>LHCBM1</i>	F	GGGATGCAGGTCTGAGCGGTTT	<i>AOT4</i>	F	TTGGCGGTGAGGTAGGAACAGACG
	R	CCTGCGCCTGTCGGGACCTT		R	TGCCGCCAGACAGGGAACCAC
<i>LHCBM2</i>	F	CGGTTGCGCGTTGTTTTCTCG			
	R	GCTTGCCTCCTTCGTGTCATTGTC			
<i>LHCBM3</i>	F	GCGTCGCGTCTGATGGAGTT			
	R	CGGGACCGCGTAAACCACAAAC			
<i>LHCBM4</i>	F	CAGGAGTGGAGTGGAGCGGTTTG			
	R	GTAAGGGCCATTTGCGGACATCA			
<i>LHCBM5</i>	F	GATGGCGCGTTTGTGATGG			
	R	TGGAGGGCTTTGTTCCGGTGTG			
<i>LHCBM6</i>	F	CAGCGGCGTATATTGGCACTTTG			
	R	CGCGAACCACGTCATCACCACCTA			

Supplemental Table 2. Primers used for qPCR analyses.

SAMPLE	SEQUENCING				ALIGNMENT							
	Sequencer	Cycles	Bases (Mb)	Total sequences	Aligned sequences: Reads & percent of total sequences		First round (%)	Recursive (%)	Uniquely aligned sequences: Reads & percent of total sequences		First round (%)	Recursive (%)
WT+S #1	GA-II	35	242	6908625	6271828	90.78	89.81	10.19	5719740	82.79	89.68	10.32
WT+S #2	GA-II	35	241	6893670	6259102	90.79	89.92	10.08	5714298	82.89	89.80	10.20
WT+S #3	GA-II	35	253	7218288	6556370	90.83	90.42	9.58	5943997	82.35	90.44	9.56
WT-S #1	GA-II	35	236	6739832	5927797	87.95	88.06	11.94	5134161	76.18	87.74	12.26
WT-S #2	GA-II	35	213	6098965	5304341	86.97	89.01	10.99	4574240	75.00	88.65	11.35
WT-S #3	GA-II	35	222	6332609	5636918	89.01	90.41	9.59	5051425	79.77	90.17	9.83
WT-S #4	GA-II	35	207	5914863	5245799	88.69	90.10	9.90	4684122	79.19	89.82	10.18
snrk2.1 +S #1	GA-I	35	148	4219945	3610635	85.56	89.54	10.46	3245094	76.90	89.48	10.52
snrk2.1 +S #2	GA-I	35	149	4263425	3642719	85.44	89.48	10.52	3258905	76.44	89.38	10.62
snrk2.1 -S #1	GA-I	35	158	4522337	3835547	84.81	81.24	18.76	3406669	75.33	81.09	18.91
snrk2.1 -S #2	GA-I	35	168	4800015	4066764	84.72	81.05	18.95	3612426	75.26	80.91	19.09
<b>Total/ Average</b>			<b>2237</b>	<b>63912574</b>	<b>56357820</b>	<b>87.78</b>	<b>88.09</b>	<b>11.91</b>	<b>50345077</b>	<b>78.37</b>	<b>87.92</b>	<b>12.08</b>

**Supplemental Table 3. Summary statistics for RNA-seq samples and sequence alignment.** Sequencing details and alignment summary statistics are provided for each individual Solexa lane, which are grouped, named and numbered after the corresponding biological sample. Alignment statistics before and after filtering to unique hits are presented in two separate blocks. Percentages of aligned sequences refer to the total number of sequences for each particular lane. Percentages for different rounds of alignment provide the specific contribution of each round to the total number of aligned sequences. On average, and regardless of the multiplicity filter, the recursive alignment constitutes 12% of the total alignment.



## Supplemental Methods

1. *RNA-seq references*: The program SOAP v.1 (Li et al., 2008) was used for alignment of processed sequences, with parameters  $r = 2$  (report all matches for a given read), a seed size optimized to the size of the read (from  $s \in [9,12]$  for reads length  $k \in [21,35]$ ) and  $v = 2$  (up to two mismatches) for raw reads ( $k = 35$ ) or  $v = 0$  (no mismatches) for  $k \in [21,33]$ . For the analysis of the quality and reproducibility of RNA-seq data, we follow the notation reported by Bullard et al., (2009), along with some of the techniques therein. For differential expression, we used the Poisson model described by Audic and Claverie (1997). Sets of “p-values” generated by this model were used to address lane-to-lane and run-to-run variations along with those genes differentially expressed between experiments. Additional references that were used to some extent for the analysis are cited in the following paragraphs. The more relevant details are explained below, and the reader is referred to the original references for further information. The more relevant details are explained below, and the reader is referred to the original references for further information.

2. *Definitions and notations*: Throughout this work, matrix subscripts are used to denote individual lanes ( $i \in [1, I]$ ), genes or sets of genomic bases ( $j \in [1, J]$ ), and aligned read lengths ( $k \in [k_{\min}, k_{\max}]$ ). The mRNA samples from 4 different experiments ( $a(i) \in [a_1, a_4]$ ) were sequenced in  $I = 11$  lanes corresponding to ( $c(i) \in [c_1, c_8]$ ) different flow cells. The set of curated genes provided in the JGI v3.1 *Chlamydomonas* genome annotation includes 14,598 gene models, with a few additional models added to produce a final set of  $J=14,605$  gene models. We used a recursive alignment strategy (see below) in which the complete read ( $k_{\max} = 35$ ) or part of the read ( $k_{\min} = 21$ ) is mapped to the reference sequence. For each alignment (the set of genomic regions to which a particular read aligns), the most 5' positions are referred to as *hits*. For normalization purposes, we keep track of the read length, defining a *k-hit* as a hit corresponding to a read of length  $k$ . Let  $H_{ik}(x)$  denote the matrix containing the number of *k-hits* for lane  $i$  at a given genomic position  $x$ . Thus,  $H_{ijk} = H_{ik}(x \in \mathbf{M}_{jk})$  would contain the number of *k-hits* per lane for a given mappable set  $\mathbf{M}_{jk}$  (i.e. gene, see below). Replacing the corresponding subscript by a “.” symbol represents totals over all lanes, genes or read

lengths. That is, the total number of hits for a given gene and lane is denoted by  $H_{ij\bullet}$ , and the total lane counts by  $H_{i\bullet\bullet}$ . When there is no ambiguity, we use an implicit notation where indices for totals are omitted, that is,  $H_{ij} = H_{ij\bullet}$ . If groups of lanes from two different experiments were pooled for a differential expression analysis,  $H_{a(1)j\bullet}$  or  $H_{a(1)j}$  would be read as the total number of hits for the mappable set  $j$  over all lanes from experiment  $a_1$ . Variables corresponding to a particular multiplicity filter (e.g. unique hits) are represented with a superscript:  $H^1_{ijk}$  indicates the hits matrix when only unique hits are considered. No superscript is used for unfiltered variables.

Let us consider that the complete set of alignments is filtered to keep only unique  $k$ -hits, yielding the matrix  $H^1_{ik}(x)$ . This matrix can be used to build genome-wide coverage graphs that we represent on a local installation of the UCSC genome browser at <http://genomes.mcdb.ucla.edu> (Kent et al., 2002). The *point unique coverage*  $C^1_{a(i)}(y)$  can be obtained as the total number of uniquely aligned reads that overlap with base  $y$  in the genome:

$$C^1_{a(i)}(y) = \sum_k H^1_{a(i)k}(x \in [y - k + 1, y])$$

It is plotted using a modified logarithmic scale ( $\log_{10}(1 + C^1_{a(i)}(y))$ ) to regularize the logarithm in regions of null coverage.

**3. Mappability:** For each  $k \in [21, 35]$ , we compute the number of copies (multiplicity) of each possible  $k$ -mer in the genome and assign the corresponding value to every genomic coordinate. This whole-genome multiplicity allows us, for each annotated gene  $j$ , to determine the set of mappable bases  $\mathbf{M}^m_{jk}$  at a given multiplicity level  $m$  (the so-called *m-mappable set*). Its cardinality  $M^m_{jk} = |\mathbf{M}^m_{jk}|$  is denominated gene  $m$ -mappable length or *m-mappability*. This way, from the perspective of short-read alignments, we consider *mappable sets* instead of *genes*, which indicate the family of mappable sets obtained for all values of  $m$  and  $k$ . These values depend on the alignment (the range of  $k$  values) and the global properties of the genome (which enter into play via the dependency of  $\mathbf{M}^m_{jk}$  on  $m$ ).  $\mathbf{M}^m_{jk}$  could be an empty set below some threshold for  $k$ . In the simplest case, the unique hits filter would then yield the  $\mathbf{M}^1_{jk}$  family of mappable sets. For the sake of

clarity, we will use the term *gene* in most of our discussions.

For transcriptome applications,  $\mathbf{M}^m_{jk}$  needs to be further modified to account for (see Supplemental Figure 6):

- Bases that overlap with neighboring gene models or belong to more than one isoform. Different conventions can be adopted to deal with these cases. The number of gene models with different isoforms included in the JGI annotation is relatively small ( $\sim 200$ ), so we decided to analyze all of them individually. Moreover, regions of overlap between two different gene models  $(j_1, j_2)$  were removed from  $\mathbf{M}^m_{j_1k}$  and  $\mathbf{M}^m_{j_2k}$  for all values of  $k$  and  $m$ .
- A small percentage of aligned reads were observed to overlap with intronic regions. Intronic sequences from non-annotated splice variants or, more likely, mRNAs that have been polyadenylated but not fully processed, are potential sources of sequences that align with introns. Based on our data, the latter seems to dominate as the number of aligned reads overlapping introns correlates with those in neighboring exon sequence (data not shown). However, in order to avoid ambiguities, we chose to filter out all of the hits that overlap with introns, i.e., those *k-hits* located in a window of size  $k-1$  downstream of the 3' end of each exon. Thus, all genomic bases from these windows were removed from  $\mathbf{M}^m_{jk}$ .

4. *Sequence alignment*: Since RNA-seq has been used in this work as a high-throughput, high-resolution technique to perform differential expression analysis of mRNA, we could in principle use the transcriptome sequences as a reference for alignment. This choice could potentially increase the number of aligned sequences by including the exon-exon junction sequences in the alignment. In fact, it is possible to estimate the amount of sequences that could be aligned using a library of exon-exons junctions (Sultan et al., 2008). Considering that the average transcript in *Chlamydomonas* has 1,580 bases and that the genes on average have 8.33 exons (Merchant et al., 2007), approximately 14% of the 35-mer reads should align to exon-exon junctions (under conditions of uniform coverage). However, according to our alignment results,  $\sim 25\%$  of the sequences map to intergenic or intronic regions. This finding indicates that the current annotation of the *Chlamydomonas* genomes is rather incomplete, preventing us from performing a reliable transcript-based alignment. Moreover, the assignment of hits at a given multiplicity level

would be ambiguous if we do not allow for alignment to non-annotated sequence. Transcriptome annotation based on deep sequencing of DNA and RNA is an active area of research, and most likely will improve the annotation of subsequent releases of the *Chlamydomonas* genome. However, in this work we do not make use of RNA-Seq data to modify the existing annotation.

Therefore, we do not perform an alignment to annotated transcripts and instead we choose to perform a hierarchical genomic-based alignment that consists of two steps:

- a. First alignment round: Raw 35-mers were aligned with a tolerance of up to 2 mismatches and no indels. Approximately 77% of reads mapped to at least one genomic location on either of the two strands, and of these, 89% mapped uniquely (Supplemental Table 3). The previous percentage compares favorably with previously reported RNA-seq studies (Marioni et al., 2008; Mortazavi et al., 2008), and is similar to additional *Chlamydomonas* RNA-seq data from on-going experiments (data not shown).
- b. Recursive alignment: Along with exon-exon junctions, there are other reasons why reads do not align to the genome (e.g. sequencing errors, adaptor sequences, poly A signals) and these result in a mismatch probability that is higher on both the 5' and 3' ends of the reads. In an attempt to align these reads, we used a trimming procedure in which we recursively trimmed one base at a time (up to  $k_{\min} = 21$ ) at either the 5' or 3' ends of the reads until we found a match (perfect) to the genome. For each step, we collected the set of non-aligned reads and build a new library that contained both the trimming of one base at the 5' or 3' end. No mismatches were allowed, and only non-ambiguous alignments (unique genomic position and unique trimming) were considered positive hits. However, reads that provided ambiguous alignments (both the 5' and 3' trimmed reads aligned to at least one genomic base, or any of the trimmed reads aligned to more than one genomic base) were discarded for the next step. The particular choice of  $k_{\min} = 21$  is related to the observed multiplicity of the *Chlamydomonas* genome, and the fact that for k-mers smaller than  $k_{\min} = 21$ , the number of genes with a mappable fraction of more than 90% decreases significantly (data not shown). Detailed summary statistics are shown in Supplemental Table 3. The trimming procedure allowed us to recover close to 50% of the reads not readily

aligned in the initial round of alignment, bringing the total fraction of mappable reads to ~88%. Moreover, we expect that expression estimates and reproducibility for genes with low expression and a high number of exons will benefit significantly from this additional alignment step. Supplemental Figure 7 shows the browser tracks for a particular region of the genome, where coverage from the recursive alignment round is compared to that of the initial alignment. It is clear that there is a significant improvement in the alignment of sequences across exon boundaries, which might be critical for the detection of small exons.

*5. Statistical analysis of RNA-seq data:* One of the main advantages of high-throughput sequencing technology applied to transcriptome analysis is its quantitative nature. As compared to microarray technology, where hybridization intensities do not correlate strongly with transcript abundance, sequencing-based technologies provide a digital random sampling of the original mRNA pool. Thus, under high-throughput conditions, the total number of hits  $H^1_{ij}$  for a particular gene can be modeled as a random variable from a Poisson distribution. Several approaches have been proposed to perform Poisson statistics on RNA-seq count data (Audic and Claverie, 1997; Marioni et al., 2008; Mortazavi et al., 2008; Bullard et al., 2009). In the present work, we applied the method presented in Audic and Claverie (1997) to address the significance of hits-per-gene variations between different lanes, runs or experiments. Let  $x$  denote the total observed number of hits associated with a given genomic region or mappable set  $\mathbf{M}^1_{jk}$  and a lane or group of lanes  $a_1$  (that is,  $x = H^1_{a_1,j}$ ). Under conditions typical of high-throughput sequencing data (the total  $N_1 = H^1_{a_1}$  of unique  $k$ -hits obtained from a random subset of cDNAs from the original mRNA population is much bigger than  $H^1_{a_1,j}$  for all  $j$ ) the probability distribution function for  $x$ ,  $\mathbf{p}(x)$ , which approximates a Poisson distribution

$$\mathbf{p}(x) \sim \frac{e^{-\lambda_1} \lambda_1^x}{x!}$$

where  $\lambda_1$  is the expected value for  $H^1_{a_1,j}$ . If one data set is compared with another from a new group of lanes  $a_2$  (such that  $N_2 = H^1_{a_2}$ ), we wish to compute the conditional probability  $\mathbf{p}(y|x)$  of having  $y = H^1_{a_2,j}$  hits for the same mappable set  $\mathbf{M}^1_{jk}$ . The expected value for  $y$  is denoted by  $\lambda_2$ . Following Audic and Claverie (1997), an

approximate closed-form solution to this problem is:

$$\mathbf{p}(y | x) = \left(\frac{N_2}{N_1}\right)^y \frac{(x+y)!}{x!y! \left(1 + \frac{N_2}{N_1}\right)^{x+y+1}}$$

given that:

- $x$  is used as the maximum likelihood estimate for  $\lambda_1$ ,
- an uninformative uniform prior in the interval  $[0, \infty)$  is assumed for the unknown distribution  $\mathbf{p}(\lambda)$ ,
- the ratio of the expected values  $(\lambda_1, \lambda_2)$  is fixed to be equal to the ratio of hits totals  $(N_1, N_2)$ .

Thus, making use of the cumulative distribution function for  $\mathbf{p}(y | x)$ :

$$\mathbf{q}(y' \leq y | x) = \sum_{y'=0}^{y' \leq y} \mathbf{p}(y' | x)$$

we can define a two-tailed p-value for differential expression of the observations  $(x, y)$  for the same gene  $j$  as  $p_j = 2 * \mathbf{q}(y' \leq y | x)$ . Taking advantage of the  $x \leftrightarrow y$  symmetry of  $\mathbf{p}(y | x)$ , if the one-sided cumulative probability is  $> 0.5$ , we first subtracted it from 1 and then multiplied this value by 2 in order to obtain the two-sided p-value.

A similar methodology can be used to address the influence of the sequencing details on hits and expression estimates. Based on the hypothesis that two samples have the same statistical properties and that technical reproducibility is perfect, the distribution of p-values should be uniform. Supplemental Figure 4 shows a set of qq-plots where the p-values of different statistical tests for hits totals  $H^1_{ij}$  ( $i \in [1, 11]$ ) are compared to the uniform distribution in  $[0, 1)$ . Plots 1 to 8 therein correspond to technical replicates of the same samples, but only 1 and 2 correspond to samples that were run in the same flow cell. The last 4 plots correspond to different samples (different stress levels or genetic background). Thus, it is readily observed that lane-to-lane and run-to-run variations are detectable. However, the fraction of genes with very small p-values is much bigger when different samples are compared. We further explore previous p-value distributions in order to estimate false discovery rates. Posterior probabilities of differential expression,  $\text{PPDE}(p) \in [0, 1)$ , were computed from the previous p-values using the mixture-models

approach implemented in (Allison et al., 2002) in order to estimate false discovery rates. Briefly, the p-value distributions are fitted to a mixture of a beta distribution (representing the set of differentially expressed genes) and a uniform distribution (non-differentially expressed). To this end, we made use of an implementation in R included in the package cyberT (Baldi et al., 2001).  $PPDE(p_j) \sim 1$  implies that gene  $j$  has a large probability of being differentially expressed. This method allows us to estimate the false discovery rates at any give p-value threshold,  $PPDE(< p_j)$ . In practice, we do not impose any particular threshold for  $PPDE(p_j)$ , but typically more than 95% percent of the target genes in each comparison obey  $PPDE(p) > 0.99$ .

To further explore the reproducibility of our experiments, we can inspect the variability of expression levels between different lanes. Following Bullar et al. (2009), we define a *normalized expression measure* as

$$E^1_{a(i)j} = \frac{c}{H^1_{a(i)}} H^1_{a(i)j}$$

for a set of lanes  $a(i)$  and a constant factor  $c$  (equal to  $1 \times 10^6$  in this work). Supplemental Figure 4 shows a set of mean-difference scatter-plots for the same pairwise comparisons as in Supplemental Figure 5. The expression fold-changes  $E^1_{a(i)j_1} / E^1_{a(i)j_2}$  are plotted as a function of the overall expression measure  $\sqrt{E^1_{a(i)j_1} / E^1_{a(i)j_2}}$  in a modified logarithmic scale  $\log_{10}(1+x)$ . Again, some lane-to-lane or run-to-run variations are observed, mostly when replicates sequenced in different flow cells are compared. However, it is readily seen that variability in expression fold-changes for the same sample are negligible relative to those from different samples.

## SUPPLEMENTAL REFERENCES

- Allison, D.B., Gadbury, G.L., Heo, M.S., Fernandez, J.R., Lee, C.K., Prolla, T.A., and Weindruch, R.** (2002). A mixture model approach for the analysis of microarray gene expression data. *Comput Stat Data An* **39**, 1-20.
- Audic, S., and Claverie, J.M.** (1997). The significance of digital gene expression profiles. *Genome Res.* **7**, 986-995.
- P. Baldi and A.D. Long** (2001). A Bayesian framework for the analysis of microarray expression data: regularized t-test and statistical inferences of gene changes. *Bioinformatics* **17**, 509-519.

- Bullard, J.H., Purdom, E.A., Hansen, K.D., Durinck, S., and Dudoit, S.** (2009). Statistical inference in mRNA-Seq: exploratory data analysis and differential expression. U.C. Berkeley Division of Biostatistics Working Paper Series **247**.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D.** (2002). The human genome browser at UCSC. *Genome Res.* **12**, 996-1006.
- Ledford, H.K., Chin, B.L., and Niyogi, K.K.** (2007). Acclimation to singlet oxygen stress in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **6**, 919-930.
- Li, R., Li, Y., Kristiansen, K., and Wang, J.** (2008). SOAP: short oligonucleotide alignment program. *Bioinformatics* **24**, 713-714.
- Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y.** (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res.* **18**, 1509-1517.
- 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., Marechal-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., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-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., Riano-Pachon, 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., Martinez, D., Ngau, W.C., Otilar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., and Grossman, A.R.** (2007). The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science* **318**, 245-250.
- Mortazavi, A., Williams, B.A., McCue, K., Schaeffer, L., and Wold, B.** (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat. Methods* **5**, 621-628.
- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., and Snyder, M.** (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320**, 1344-1349.
- Sultan, M., Schulz, M.H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., Schmidt, D., O'Keefe, S., Haas, S., Vingron, M., Lehrach, H., and Yaspo, M.L.** (2008). A

global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321**, 956-960.

