Plant Biotechnology Journal

Plant Biotechnology Journal (2013) **11**, pp. 717–733 doi: 10.1111/pbi.12062

New insights into *Chlamydomonas reinhardtii* hydrogen production processes by combined microarray/RNA-seq transcriptomics

Jörg Toepel[1], Maike Illmer-Kephalides[1], Sebastian Jaenicke[2], Jasmin Straube[2], Patrick May[3], Alexander Goesmann[2] and Olaf Kruse[1,*]

[1]*Algae Biotechnology & Bioenergy Group, Department of Biology/Center for Biotechnology, Bielefeld University, Bielefeld, Germany*

[2]*Computational Genomics, CeBiTec,Bielefeld University, Bielefeld, Germany*

[3]*AG Bioinformatics, Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm, Germany*

Received 11 October 2012;
revised 7 January 2013;
accepted 9 February 2013.
**Correspondence* (fax +49-521-10612290; email Olaf.kruse@uni-bielefeld.de)

Keywords: transcriptomics, hydrogen production, *Chlamydomonas reinhardtii*.

Summary

Hydrogen production with *Chlamydomonas reinhardtii* induced by sulphur starvation is a multiphase process while the cell internal metabolism is completely remodelled. The first cellular response is characterized by induction of genes with regulatory functions, followed by a total remodelling of the metabolism to provide reduction equivalents for cellular processes. We were able to characterize all major processes that provide energy and reduction equivalents during hydrogen production. Furthermore, *C. reinhardtii* showed a strong transcript increase for gene models responsible for stress response and detoxification of oxygen radicals. Finally, we were able to determine potential bottlenecks and target genes for manipulation to increase hydrogen production or to prolong the hydrogen production phase. The investigation of transcriptomic changes during the time course of hydrogen production in *C. reinhardtii* with microarrays and RNA-seq revealed new insights into the regulation and remodelling of the cell internal metabolism. Both methods showed a good correlation. The microarray platform can be used as a reliable standard tool for routine gene expression analysis. RNA-seq additionally allowed a detailed time-dependent study of gene expression and determination of new genes involved in the hydrogen production process.

Introduction

Research in hydrogen production in *Chlamydomonas reinhardtii* is of specific interest, because alternative renewable energy sources are highly desired and needed in the near future. It is known that *C. reinhardtii* produces molecular hydrogen under anaerobic conditions, a phenomenon that can be induced by sulphur deprivation (Melis and Happe, 2001). Under such conditions, *C. reinhardtii* cells remodel their internal metabolism and use hydrogen production as a valve system to prevent over-reduction of the chloroplast stroma. During the last decade, intensive research has been conducted on aspects of the cellular metabolism connected to the H_2 production process. These investigations include cellular processes such as fermentation (Catalanotti *et al.*, 2012; Hemschemeier *et al.*, 2008b; Mus *et al.*, 2007; Philipps *et al.*, 2011), photosynthesis and respiration (Chochois *et al.*, 2009, 2010; Desplats *et al.*, 2009; Esper *et al.*, 2006; Kruse *et al.*, 2005; Lecler *et al.*, 2011; Tolleter *et al.*, 2011; Torzillo *et al.*, 2009) and the regulation of the hydrogenase HydA (Happe and Kaminski, 2002; Happe *et al.*, 2002). Several new findings have demonstrated remarkable reorganization capacities of *C. reinhardtii*, especially the compensation of knocked-out or down-regulated genes related to fermentation processes (Catalanotti *et al.*, 2012; Grossman *et al.*, 2011; Magneschi *et al.*, 2012; Philipps *et al.*, 2011). Additionally, hydrogen production has been investigated through several systems biology approaches regarding changes in the metabolome (Doebbe *et al.*, 2010; Matthew *et al.*, 2009), in the proteome (Chen *et al.*, 2010) and in the transcriptome (Mus *et al.*, 2007; Nguyen *et al.*, 2008) of *C. reinhardtii* to understand cellular adaptation in detail and to identify new targets or bottlenecks to improve hydrogen yield. A particular research focus was set on the cellular adaptation to sulphur stress induced anaerobiosis, resulting in the induction of distinct genes responsible for sulphate transport and assimilation, accompanied by a repression of gene expression of the vast majority of genes related to photosynthetic processes (Mus *et al.*, 2007; Nguyen *et al.*, 2011). These gene adaptation mechanisms are reported to be followed by the over-expression of gene products required for starch and lipid synthesis (Doebbe *et al.*, 2010; Matthew *et al.*, 2009) and in acclimatization/modulation processes that include changes in the amino acid composition of certain target proteins (Doebbe *et al.*, 2010). The hydrogen production phase is characterized by the expressed and active hydrogenase, which is competing under anaerobic conditions for electrons derived from photosynthesis (Chochois *et al.*, 2009, 2010; Doebbe *et al.*, 2010) with fermentative-related enzymes, such as alcohol dehydrogenase (Hemschemeier and Happe, 2011; Hemschemeier *et al.*, 2008a,b; Philipps *et al.*, 2011).

To improve hydrogen production in *C. reinhardtii*, specific knowledge of the bottlenecks in all pathways and processes is required. Therefore, new, detailed analyses of the transcriptome and proteome under different conditions are necessary to elucidate these limitations and identify potential targets for improvement. Rapid progress in genome annotation of *C. rein-*

hardtii resulted in the design of a new microarray platform for advanced and general transcriptome analyses (Toepel *et al.*, 2011). *Chlamydomonas reinhardtii* full-genome microarrays enable us to determine expression level variations of ~11 000 gene models. New annotations, however, predicted nearly 20 000 gene models, and RNA-seq data have been used to provide new insights into transcriptomic changes during nitrogen (Miller *et al.*, 2010), sulphur (Boyle *et al.*, 2012; Gonzalez-Ballester *et al.*, 2010) or carbon dioxide (Fang *et al.*, 2012) limitation. In contrast to RNA-seq, microarrays are relatively inexpensive, reliable systems for use on routine basis and have the potential to give a rapid overview of variations in transcript levels. RNA-seq platforms have many advantages in comparison with microarray technologies, such as higher gene coverage and increased sensitivity for differential gene expression. Additionally, predictions of new gene models and splicing variations can be realized by RNA-seq as well as the detection and characterization of mutation sites (Smith *et al.*, 2008). However, data obtained from RNA-seq require stringent examination, and reproducibility of results is often low. As a typical consequence, an overestimation of highly abundant genes and a length-dependent amplification have been reported by this method. Full data analysis and data normalization for RNA-seq experiments is not yet standardized (Liu *et al.*, 2011; Malone and Oliver, 2011; Oshlack *et al.*, 2010; Wang *et al.*, 2011). Here, we compare expression levels of induced/repressed transcripts generated by both methods to further characterize this process and identify new potential bottlenecks of hydrogen production in *C. reinhardtii*. However, full-data analysis and data normalization is not standardized yet (Ramskold *et al.*, 2012). To validate RNA-seq data and to confirm the usability of microarrays, we intended to compare expression level of induced/repressed transcripts of *C. reinhardtii* during hydrogen production with both methods to further characterize this process and to identify new potential targets to reduce limitations by bottlenecks.

Results

Hydrogen production in a mutant defective in non-photochemical quenching

Within this project, the *C. reinhardtii* strain *npq*4 (Niyogi *et al.*, 1997) and its corresponding wild-type 4A+ were used to identify key elements of hydrogen production processes. *Npq*4 was used as a reference strain for these comparative analyses to evaluate the accuracy of both RNA-seq and microarray analyses. *Npq*4 is a mutant defective in non-photochemical quenching, which was suggested to yield in clear differences regarding its transcriptome upon stress induction by sulphur depletion and anaerobiosis (Bonente *et al.*, 2011; Peers *et al.*, 2009). Hydrogen production from *C. reinhardtii* induced by anaerobiosis via sulphur depletion is a multiphase process, and *Npq4* demonstrated a typical course of hydrogen production; a lag phase of 24 h until hydrogen production began, followed by a ~72-h production phase with a total production of around 120–150 mL H_2 per litre culture (Figure 1). This total hydrogen production rate in the mutant was only marginally higher when compared to the corresponding wild-type *4A*+ (90–110 mL H_2). PAM fluorescence measurements revealed that effective photosynthetic quantum yields decreased in both strains from 0.5 to 0.05 within the first 24 h followed by consistent photosynthetic efficiency reflected by quantum yields of ~0.01 for the remaining duration of the experiment (Figure 1).

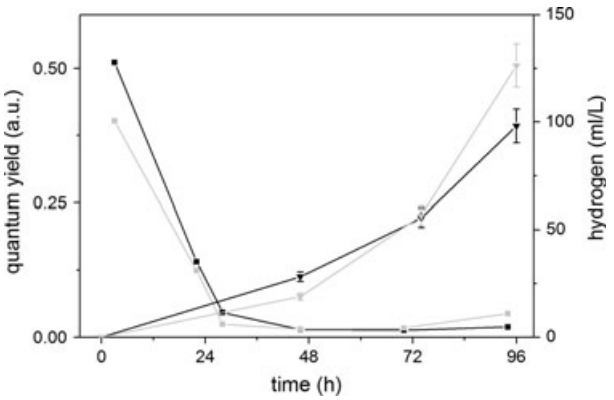

Figure 1 Hydrogen production (▼) and effective quantum yield (■) during hydrogen production of *Chlamydomonas reinhardtii* strains *npq*4 (grey) and *4A*+ (black).

Comparison of microarray and RNA-seq to identify differential gene expression

Microarray and RNA-seq analyses were performed to identify and determine changes of transcript levels for distinct genes over the entire period of hydrogen production in *npq*4. The subsequent analysis of microarray data included the expression pattern for ~10 000 gene models (Toepel *et al.*, 2011). We used a gene expression threshold with a cut-off of twofold for both up- or down-regulation and could identify 603 gene models with a down-regulation and 635 gene models with an up-regulation (see Figure 2 solid box: 635 genes red; 603 genes green). Quantitative RT-qPCR of several control genes were in accordance with our microarray data (see Supporting information) and are also in agreement with previous studies (Nguyen *et al.*, 2011) performed with a different first-generation microarray platform (Eberhard *et al.*, 2006). It should be noted that the microarray platform used in our study included almost 4000 new gene models including a large number of unknown gene models (Toepel *et al.*, 2011). As a consequence, the majority of differentially expressed genes observed in this study are unknown or not yet fully

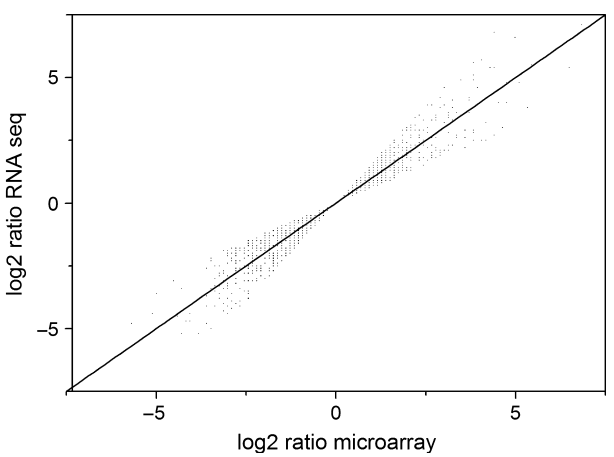

Figure 2 Overview of differential expressed gene models in *Chlamydomonas reinhardtii* during hydrogen production determined with microarray (solid line; total: 1238) and RNA-seq (dotted line: total: 1598). Red dots represent up-regulated genes and green dots down-regulated genes.

characterized. However, we could confirm up-regulation of most of the genes identified in a previous study (Nguyen *et al.*, 2011), for example, genes coding for proteins involved in sulphur metabolism/catabolism as well as transcripts related to lipid and starch metabolism (see Table 1). Additionally, the up-regulation of genes related to the pentose phosphate cycle and for genes related to fermentation processes could be confirmed. One major result, the gene expression of the isocitrate lyase (*Icl*), the key enzyme for the TCA cycle shunt, could be confirmed with our experiments. We were also able to verify the down-regulation of the majority of photosynthetic genes (PS I; PS II and ATPase) and increased transcript levels of *Lhcbm*9 and *Lhcsr*1 (Table 1), two distinct genes related to light harvesting and energy quenching (determined with RT-PCR from (Nguyen *et al.*, 2011). Additionally, as we used *Npq*4, a mutant deficient in *Lhcsr*3.1 and *Lhcsr*3.2, we could control the specificity of the read alignment. Within both methods, RNA-seq and microarray analyses, no *Lhcsr*3.1 and *Lhcsr*3.2 reads could be identified in data sets derived from *npq*4 cells during hydrogen production. In contrast, comparative wild-type analyses showed a strong gene expression of *Lhcsr*3 genes, clearly demonstrating the reliability of both methods (Nguyen *et al.*, 2011; Toepel *et al.*, 2011).

In a second step, we used the microarray data for comparison with RNA-seq-determined transcript abundance levels. RNA-seq analysis demonstrated at least 52 million reads per sample with around 18 million unique reads. The *Chlamydomonas* genome (phytozome 4.3) was applied for mapping, and we calculated the gene expression for ~13 000 gene models (based on an estimation of total ~18 000 gene models), which equates coverage of ~70% with a correlation factor of 0.8–0.88 between all samples (except the control). Estimations of *log*2-fold changes out of the RPKM data were performed. For our analysis, we included only differentially expressed genes with an average twofold down-regulation (730) or up-regulation (868) compared to the control. Intriguingly for 260 gene models we identified reads only in the control sample, 15 of these gene models showed high transcript level just detectable under normal growth conditions, for example, *Psa*O (*Cre07.g334550*), *Cah*1 (*Cre04.g223100*), one triose phosphate transporter (*Tpt*2; *Cre06.g263850*), one porphobilinogen deaminase (*Cre16.g663900*), one RNA polymerase (*Cre02.g086750*) and several unknown genes (see Supporting information). In contrast, 280 newly induced genes during hydrogen production were counted, including many sulphur-related genes, stress-related genes, cytochromes and many unknown genes.

Comparison between both transcriptomic data sets was performed to determine whether the two independent techniques provided similar gene expression data and whether both methods display the same transcript expression pattern over the period of time. In Figure 3, a direct comparison for all differentially expressed genes (with *log*2-fold changes) derived from microarray analysis and from RNA-seq was plotted. We found overall the same tendency in transcript abundance for 1290 gene models and for 786, an identical gene regulation above the threshold level. There is a good correlation between both data sets, however, with differences in the dynamic range of changes in expression level (postulated previously by Gonzalez-Ballester *et al.*, 2010). This enabled us to gain more information from RNA-seq data regarding time-dependent gene expression during the experiments.

As mentioned above, ~63% (786 of 1238) of twofold differentially expressed genes derived from microarray analysis could be confirmed by identical expression profiles when RNA-seq was applied (see Figure 2). From these 786 gene models, 395 genes were down-regulated whereas 391 genes showed an up-regulation in both methods. Overall, ~45% of the differential expressed genes (320 of 786 genes) are of unknown function. The remaining gene models (55%) are related to lipid and starch metabolism (27 genes), sulphur metabolism (ten genes), photosynthesis (14 genes) and general stress response (five genes). In addition, numerous genes associated with flagella assembly and cell cycles were identified as being mainly down-regulated.

Time-resolved gene expression analysis during hydrogen production

To improve the time-resolved analysis of our transcript data, we used expression levels of all transcripts and clustered all genes according to their expression pattern. Clustering of differential expressed genes resulted in formation of distinct groups. In Figures 4 and 5, the clusters for both methods are presented. Microarray data suggested that the majority of genes are highly induced or repressed already 24 h after sulphur deprivation with no further changes during the rest of the experiment. In contrast, RNA-seq data seemed to be more sensitive, as we could distinguish between genes induced after 48 h, but also a considerable amount of genes where the expression profile changed in the later phase of the experiment. Of particular note was that we succeeded to identify a group of genes which showed identical transcript expression profile in both methods.

Strongly up-regulated genes were found in RNA-seq clusters A-D (446 genes), and 629 genes from microarray data are shown in clusters C and D. In such clusters, early and consistent up-regulated genes were found including sulphur-related genes (21 gene models, e.g., *Sua*1, *Slt*1, *Sir*1, *Ecp61* (*Cre09.g409300*), *Ecp*76 and *Ecp*88) and stress response genes (5 gene models), reductases (seven), genes encoding peptidases (five) and several cytochromes (five). With *Lhcbm*9 and *Lhcsr*1 and one starch phosphorylase enzyme (*Cre07.g336950*), three photosynthesis-related genes were also detected. An increased gene expression was determined for one chlorophyllase and some other chlorophyll-degrading genes.

In clusters H and I (RNA-seq) and clusters F and G (microarray), we identified genes that were strongly down-regulated including the majority of photosynthetic genes, for example, both small RUBISCO subunits (*Rbcs*1 and 2), *Psb*O, *Psb*Q, *Psa*G, *Psa*H, *Psa*K, *Psa*E, *Psa*L, *Lhca*1, *Lhca*5, *Lhca*7, *Lhcbm*1, *Lhcbm*5 and *Lhcbm*7. Other photosynthetic genes also demonstrated a decreased gene expression, like plastocyanin (*Pet*E, *Pcy*1; *Cre03.g182551*) and several components of the cytochrome b_6f complex (*Pet*C, *Pet*F, *Pet*M, *Pet*N) also showed a dramatic decreased gene expression. Just a few transcripts could be aligned for each subunit of the complex.

Specific early up-regulated genes were found by RNA-seq (clusters A, B and C); interestingly, this group includes enzymes like the mitochondrial pyruvate dehydrogenase kinase (*Pdk*3; *Cre05.g241750*), inhibitor of the citrate cycle starting enzyme), while the pyruvate dehydrogenase was down-regulated much later.

Differential gene expression of sulphur deprivation-related gene models

It is known that *C. reinhardtii* cells respond first to the lack of sulphur by increased sulphur assimilation (see Table 1 for related genes) and redistribution of internal sulphur. In our experiments,

Table 1 Differential gene expression for *Chlamydomonas reinhardtii* grown under hydrogen-producing conditions, determined with microarray and RNA-seq. Shown are the log2 ratios for both methods as mean values over the time course of the experiments

Locus	Name	Description	Log2 ratio microarray	Log 2 ratio RNA seq
Sulphur-related gene models				
Cre16.g671400	ars1	Arylsulfatase	4.3	6.6
Cre16.g671350	ars2	Arylsulfatase	2.0	2.4
Cre01.g012100	ars4	Arylsulfatase	0.3	1.3
Cre10.g431800	ars5	Arylsulfatase	−3.1	−4.8
Cre02.g107450	ATS2	ATP-sulfurylase	3.9	4
Cre03.g203850	ATS1	ATP-sulfurylase	0.7	2.9
Cre12.g517150	APR3	APS reductase 3	1.7	1.9
Cre06.g273750	SUA1	Chloroplast sulphate transporter	2.7	2.1
Cre03.g160400	SAC1	Sulphur acclimation protein	−0.6	−1.7
Cre01.g021200	CGL47	Control protein	0.4	1.8
Cre01.g012150	PMSR3	Peptide methionine sulphoxide reductase 3	1.6	1.7
Cre06.g257650	PMSR4	Peptide methionine sulphoxide reductase 4	2.4	5.9
Cre03.g210200	SOX	Sulphite oxidase	2.2	3.0
Cre05.g232800	SRX	Sulfiredoxin	−1.4	−2.9
Cre16.g693150	SIR1	Sulphite reductase	3.1	6.3
Cre08.g365700	SIR2	Sulphite reductase	0.6	1.9
Cre03.g210200	SOX	Sulphite oxidase	2.2	3
Cre12.g502600	SLT1	Sulphate transporter	5.6	5.4
Cre10.g445000	SLT2	Sulphate transporter	3.8	6.8
Cre02.g138950	SULTR4;1	Sulphate transporter 4.1	−2.6	−3.4
Cre13.g573250	STR16	Sulphate transporter	−2.3	−5.6
Cre13.g607050	STR1	Sulphate transporter	0.7	4.4
Cre13.g597450		Sulphate transporter (glutaredoxin)	1.3	2.9
Cre16.g656400	SQD1	Sulfoquinovosyldiacylglycerol 1	0.8	3.3
Cre16.g689150	SQD2	Sulfoquinovosyldiacylglycerol 2	0.0	n.d
Cre01.g038550	SQD2	Sulfoquinovosyldiacylglycerol 2	1.8	1.7
Cre14.g615000	MSRB2	Sulphate transporter	0.2	3.4
Cre02.g097900	AAT5	Aspartate aminotransferase 5	1.6	2.7
Cre07.g319400	ACD1	D-cysteine desulfhydrase	0.2	3.2
Cre01.g012150	ATMSRA3	Peptide methionine sulphoxide reductase 3	1.6	1.7
Cre01.g036750		S-adenosyl-L-methionine-dependent methyltransferase	1.8	3
Cre06.g288550	ECP 76	Extracelullar protein	3.2	7.9
Cre12.g556000	ECP 88	Extracelullar protein	3.2	7.9
Photosynthetic-related gene models				
Cre06.g284200	Lhcbm9	PS II light harvesting complex	5.3	7.6
Cre03.g156900	Lhcbm5	PS II light harvesting complex	−0.3	−5.6
Cre06.g283950	Lhcbm4	PS II light harvesting complex	−1.4	−3.3
Cre12.g548400	Lhcbm2	PS II light harvesting complex	1.1	0.1
Cre12.g548950	Lhcbm7	PS II light harvesting complex	−3.9	−3.1
Cre23.g766250	Lhcbm1	PS II light harvesting complex	−1.3	−2.3
Cre56.g791050	PSAD-2	PS I subunit D-2	−1.7	−3.2
Cre10.g420350	PSAE-2	PS I subunit E-2	−1.6	−4.0
Cre09.g412100	PSAF	PS I subunit F	−1.5	−2.1
Cre12.g560950	PSAG	PS I subunit G	−1.5	−4.3
Cre07.g330250	PSAH-1	PS I subunit H-1	−1.5	−5.7
Cre17.g724300	PSAK	PS I subunit K	−1.3	−6.4
Cre12.g486300	PSAL	PS I subunit l	−2.0	−4.5
Cre02.g082500	PSAN	PS I reaction center PSI-N	−1.3	−3.1
Cre27.g775100	PSAP	PS I subunit P	4.1	4.8
Cre07.g334550	PSAO	PS I subunit P	n.d.	−10.0
Cre02.g124700	MDB2	Nac2 factor	n.d.	−2.8
Cre05.g243800	PSB27	PS II family protein	−0.6	−0.6
Cre08.g372450	PSBQ	PS II subunit Q	−0.6	−3.3
Cre06.g261000	PSBR	PS II subunit R	1.4	1.3

Table 1 Continued

Locus	Name	Description	Log2 ratio microarray	Log 2 ratio RNA seq
Cre11.g475250	PSBW	PS II reaction center W	1.2	−0.4
Cre02.g082750	PSBX	PS II subunit X	−1.8	−5.2
Cre02.g132800	PSBO1	PS II oxygen-evolving complex 1	−1.5	−5.0
Cre16.g650100	PETN	Cytochrome b6f complex PetN	−2.7	−6.6
Cre18.g744400	PETC	Cytochrome b6f complex PetC	−2.7	−4.3
Cre03.g182551	PETE1	Electron transporter; plastocyanin 1	−3.0	−5.0
Cre14.g626700	PETF	Cytochrome b6f complex PetF	−1.3	−7.5
Cre12.g546150	PETM	Cytochrome b6f complex PetM	−2.1	−6.2
Cre06.g283050	LHCA1	PS I light harvesting complex	−2.6	−2.9
Cre12.g508750	LHCA6	PS I light harvesting complex	−2.7	−3.4
Cre10.g454750	LHCA3	PS I light harvesting complex	0.8	1.2
Cre18.g749750	LHCA3	PS I light harvesting complex	−2.3	−4.8
Cre06.g272650	LHCA5	PS I light harvesting complex	−2.4	−1.5
Cre07.g344950	LHCA5	PS I light harvesting complex	−2.4	−4.0
Cre16.g687900	LHCA5	PS I light harvesting complex	−2.8	−5.2
Cre10.g425900	LHCA4	PS I light harvesting complex	−2.9	−3.0
Cre13.g598900	LHCA4	PS I light harvesting complex	−2.3	−3.9
Cre08.g365900	LHCSR1	Chlorophyll A/B binding protein 1	6.9	8.2
Cre08.g367400	LHCSR3	Chlorophyll A/B binding protein 3	0.0	0.0
Cre08.g367500	LHCSR2	Chlorophyll A/B binding protein 2	0.1	−0.1
Cre03.g148750	CLH1	Chlorophyllase 1	2.5	1.3
Cre10.g423500	HO3	Haeme oxygenase 3	0.5	2.7
Cre13.g600650		Pheophorbide a oxygenase	n.d.	2.4
Cre02.g120100	RBCS1A	RUBISCO small chain 1A	−2.4	−2.4
Cre02.g120150	RBCS2	RUBISCO small chain 2	−2.5	−3.6
Cre27.g774300		RUBISCO methyltransferase	−2.0	−4.2
Cre02.g129750		RUBISCO methyltransferase	2.2	n.d.
Cre03.g186450		RUBISCO methyltransferase	−0.2	−2.4
Cre04.g229300	RCA	RUBISCO activase	−1.3	−0.7
Cre08.g368700		RUBISCO methyltransferase	0.2	−0.8
Cre12.g503800		RUBISCO methyltransferase	−0.5	−2.3
Cre16.g661350	RMT1	RUBISCO large subunit methyltransferase	2.6	n.d.
Carbohydrate-related gene models				
Cre08.g385500	AMA1	Alpha-amylase 1	n.d.	1.6
Cre08.g362450	AMA2	Alpha-amylase 2	1.0	1.7
Cre19.g755050	ISA3	Isoamylase 3	n.d.	1.4
Cre03.g185250	SS2	Starch synthase 2	n.d.	2.4
Cre17.g721500		Granula bound starch synthase	n.d.	2.6
Cre10.g444700	SBE2.2	Starch branching enzyme 2.2	−0.7	−1.5
Cre11.g476650		Starch debranching enzyme	n.d.	1.7
Cre07.g336950	PHS	Starch phosphorylase	1.6	2.2
Cre12.g552200	PHS	Starch phosphorylase	0.4	1.9
Cre03.g175400	PGI1	Glucose-6-phosphate isomerase	0.4	1.3
Cre12.g553250	PFK5	Phosphofructokinase 5	2.2	2.8
Cre01.g029300	TPI1	Triosephosphate isomerase 1	1.4	3.5
Cre12.g485150	GAPCP-1	Glyceraldehyde-3-phosphate dehydrogenase 1P	3.9	4.5
Cre22.g763250	PGk1	Phosphoglycerate kinase	n.d.	2.0
Cre06.g272050	PGM 1	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	1.4	2.7
Cre01.g057900	PYK 3	Pyruvate kinase 3	0.4	2.3
Cre12.g533550	PYK1	Pyruvate kinase 1	−1.0	−3.1
Cre02.g141400	PCK1	Phosphoenolpyruvate carboxykinase 1	−1.4	3.0
Cre05.g241750	PDK	Pyruvate dehydrogenase kinase	0.8	2.0
Cre06.g282800	ICL	Isocitrate lyase	n.d.	1.9
Cre08.g378150	G6PD3	Glucose-6-phosphate dehydrogenase 3	3.2	4.2
Cre12.g526800	GND1	6-phosphogluconate dehydrogenase	2.3	4.9

Table 1 Continued

Locus	Name	Description	Log2 ratio microarray	Log 2 ratio RNA seq
Fermentation-related gene models				
Cre20.g758200	ADH1	Alcohol dehydrogenase	2.2	5.3
Cre01.g044800	PFL1	Pyruvate formate lyase	−2.0	0.5
Cre09.g396700	ACK1	Acetate kinase	0.6	2.7
Cre01.g057800	MLS	Malate synthase	n.d.	3.5
Cre11.g473950	PFR	Pyruvate ferredoxin reductase	n.d.	1.7
Cre10.g423250	MDH2	Malate dehydrogenase 2	n.d.	4.0
Cre03.g199800	HYDA1	Ferredoxin hydrogenase	n.d.	0.5
Cre09.g396600	HYDA2	Ferredoxin hydrogenase	1.0	3.5
Cre06.g296700		Hydrogenase assembly factor	n.d.	2.1
Cre06.g296750	HYDEF	Hydrogenase assembly factor	n.d.	3.0
Lipid metabolism-related gene models				
Cre17.g711150	FAD2	Fatty acid desaturase 2	3.0	1.7
Cre23.g765700	ACS	Acetyl-CoA synthetase	2.3	1.9
Cre05.g248150		Phospholipid/glycerol acyltransferase	0.1	n.d.
Cre09.g392300		Acyl-CoA N-acyltransferase	1.1	0.4
Cre07.g312400	DGK1	Diacylglycerol kinase1	2.7	n.d.
Cre10.g422850		Lipase	0.3	2.8
Cre02.g127300		Lipase	3.1	1.9
Cre07.g322900		Lipase	−0.1	1.6
Cre02.g126050		Lipase	−4.3	−3.6
Cre02.g121200	DGTT2	Diacylglycerol acyltransferase	n.d.	2.5
Cre06.g299050	DGTT3	Diacylglycerol acyltransferase	−0.3	−1.6
Cre02.g106400	PDAT1	Phospholipid:diacylglycerol acyltransferase	0.2	0.5
Cre07.g325550	DGK4	Diacylglycerol kinase 4	0.5	n.d.
Cre17.g707300		Phospholipid/glycerol acyltransferase	−0.8	−1.3
Cre06.g268200	TGD1	Trigalactosyldiacylglycerol 1	1.6	3.8
Cre16.g694400	TGD2	Trigalactosyldiacylglycerol 2	n.d.	1.9
Proteases and protein kinases				
Cre10.g459650		Ubiquitin-protein ligase	−1.5	−2.7
Cre12.g546650	UBC7	Ubiquitin carrier protein 7	−1.9	−1.8
Cre12.g521450	NCLPP7	Nuclear-encoded CLP protease P7	3.5	2.5
Cre10.g432150		Protein kinase	−2.8	−6.2
Cre12.g505250	CPK24	Calcium-dependent protein kinase 24	−1.4	−3.2
Cre10.g459650	KEG	Protein kinase; ubiquitin-protein ligase	−1.5	−2.7
Cre09.g413400		Protein kinase	−2.6	−2.4
Cre33.g782700		Protein kinase	−3.0	−2.2
Cre01.g001200		Protein kinase	−2.2	−1.9
Cre17.g698550		Protein kinase	−1.4	−1.9
Cre10.g457700	CPK2	Kinase cdpk isoform 2	−1.6	−1.7
Cre12.g549750	CKL2	Casein kinase-like 2	−1.6	−1.7
Cre03.g201900		Phosphatidylinositol-4-phosphate 5-kinase	−2.9	−1.6
Cre10.g466650	CPK20	Kinase 20	−3.0	−1.6
Cre16.g654300		Nucleoside diphosphate kinase	−1.8	−1.5
Cre02.g092150		Protein kinase	1.6	1.5
Cre04.g223200	MPK9	MAP kinase 9	1.6	1.9
Cre02.g126650		Protein kinase	2.0	2.7
Cre03.g173800	ATSOS4	Carbohydrate kinase	1.6	2.7
Cre06.g255350		Hydroxyethylthiazole kinase	2.0	6.1
Stress related gene models				
Cre02.g077100	GSH1	Glutamate-cysteine ligase	0.3	0.4
Cre17.g708800	GSH2	Glutathione synthetase 2	2.8	2.3
Cre03.g154950		Glutathione transferase	1.0	3.0
Cre16.g688550	GST1	Glutathione transferase	1.5	6.7
Cre12.g559800		Glutathione transferase	1.6	1.6
Cre01.g064400		Glutathione transferase	1.1	4.9

Table 1 Continued

Locus	Name	Description	Log2 ratio microarray	Log 2 ratio RNA seq
Cre10.g458450	GPX5	Glutathione peroxidase 5	0.7	2.3
Cre02.g078300	GPX6	Glutathione peroxidase 6	1.1	2.9
Cre02.g139650	APX3	Ascorbate peroxidase 3	1.3	1.6
Cre05.g233900	APX4	Ascorbate peroxidase 4	0.4	0.4
Cre01.g045700	APX5	Ascorbate peroxidase 5	0.5	2.4
Cre06.g285150	APX6	Ascorbate peroxidase 6	1.0	0.5
Cre02.g087700	SAPX	Stromal ascorbate peroxidase	2.1	−1.6
Cre17.g712100	MDAR1	Monodehydroascorbate reductase 1	−0.7	−1.7
Cre06.g271200	MDAR4	Monodehydroascorbate reductase 4	−1.5	−3.2
Cre10.g456750	DHAR2	Dehydroascorbate reductase 2	0.3	0.3
Cre01.g044700	DHAR3	Dehydroascorbate reductase 3	0	1.8
Cre10.g456050	FQR1	Flavodoxin-like quinone reductase 1	1.3	2.2
Cre10.g456000	FQR2	Flavodoxin-like quinone reductase 2	2.1	1.8
Cre10.g456100	FQR3	Flavodoxin-like quinone reductase 3	3.3	2.5
Cre03.g167150		Flavin-binding monooxygenase	2.6	2.1
Cre10.g466700		2-oxoglutarate (2OG) and Fe(II)-dependent oxygenase	3.2	5.9
Cre01.g053000		NAD-dependent glycerol-3-phosphate dehydrogenase	−1.4	−2.4
Cre10.g421700		NAD-dependent glycerol-3-phosphate dehydrogenase	0.5	2.9
Cre11.g472700	PGPS2	Phosphatidylglycerolphosphate synthase 2	2.9	0.3
Cre07.g346800		FAD-dependent oxidoreductase	2.7	3.6
Cre12.g493500		FAD-dependent oxidoreductase	2.7	2.3
Cre16.g671450		FAD-dependent oxidoreductase	3.7	n.d.
Cre02.g139200		FAD/NAD(P)-oxidoreductase	1.8	n.d.
Cre09.g395950	AOX1	Alternative oxidase	n.d.	1.4
Cre07.g350750	PTO1	Alternative oxidase	1.2	2.2
Cre03.g172500	PTO2	Alternative oxidase	1.7	0.9
Cre09.g417150	CAT2	Catalase 2	n.d.	1.6
Cre02.g096150	MSD1	Manganese superoxide dismutase 1	1.2	1.5
Cre10.g436050	FSD1	Fe superoxide dismutase 1	n.d.	2.8
Cre11.g477200		Isoflavone reductase like protein	4.9	7.7
Cre07.g355500		Oxidoreductase	3.6	2.5
Cre01.g057750		Thioredoxin	3.2	4.1
Cre07.g315100		Thioredoxin	2.4	3.4
Gene of undefined function				
Cre02.g094250		Mitochondrial substrate carrier	−3.7	−1.0
Cre28.g776600		Mitochondrial substrate carrier	n.d.	−2.6
Cre16.g650800	TIM13	Mitochondrial translocase 13	n.d.	−2.4
Cre06.g278750		Mitochondrial substrate carrier	n.d.	−2.6
Cre01.g063200	ACP1	Acyl carrier protein 1	1.0	1.6
Cre13.g577100	ACP1	Acyl carrier protein 1	−0.9	−2.5
Cre14.g621650		S-malonyltransferase	−3.1	−3.7
Cre02.g144800	NAGS2	N-acetyl-L-glutamate synthase 2	−0.1	−2.4
Cre10.g434800		Stress-inducible protein	2.6	1.7
Cre12.g495850		D-beta-Hydroxybutyrat–Dehydrogenase	n.d.	2.5
Cre07.g343050			3.7	5.4
Cre11.g471864			n.d.	10.7
Cre56.g791150			3.4	4.6
Cre03.g177250			n.d.	7.3
Cre07.g349350			1.7	2.9
Cre03.g192350			n.d.	10.7
Cre02.g113400		Unknown conserved protein	−3.1	−2.6
Cre17.g741850		RNA binding protein	−1.8	−3.9
Cre02.g114600	2-Cys Prx B	2-cysteine peroxiredoxin B	−1.4	−1.6
Cre02.g085300			−1.3	−4.3
Cre02.g087250			n.d.	−2.3
Cre03.g153450			n.d.	−1.5

Table 1 Continued

Locus	Name	Description	Log2 ratio microarray	Log 2 ratio RNA seq
Cre13.g570850			−2.4	−4.0
Cre07.g352850			−0.4	−2.1
Cre03.g164000	TEF7		−3.2	−6.0
Nutrient transport-related gene models				
Cre02.g111050	AMT1;3	Ammonium transporter 1;3	−0.4	−3.2
Cre06.g284150	AMT1;3	Ammonium transporter 1;3	n.d.	−1.8
Cre26.g773300	PHT2;1	Phosphate transporter 2;1	−1.8	−3.5
Cre02.g144750	PHT2;1	Phosphate transporter 2;1	−2.9	−3.3
Cre26.g773350	PHT2;1	Phosphate transporter 2;1	−4.4	−2.6
Cre16.g686750	PHT6	Phosphate transporter 1;6	−4.9	−4.8
Cre16.g655200	PHT2;1	Phosphate transporter 2;1	−1.0	−2.1
Cytochromes				
Cre04.g221700		Cytochrome c oxidase, subunit III	0.6	4.2
Cre03.g154350		Cytochrome oxidase 2	2.0	2.9
Cre16.g651050		Cytochrome c	1.9	6.2
Cre01.g049500		Cytochrome oxidase 2	1.8	2.2
Cre06.g304350	COX6B	Cytochrome C oxidase 6B	1.1	3.6
Cre05.g232850	COX17	Cytochrome c oxidase 17	n.d.	2.0
Cre12.g516350	COX10	Cytochrome c oxidase 10	n.d	1.8
Cre01.g051900		Ubiquinol-cytochrome C reductase iron-sulphur subunit	0.9	1.3
Cre01.g055550		Cytochrome c oxidase assembly protein	n.d.	2.8

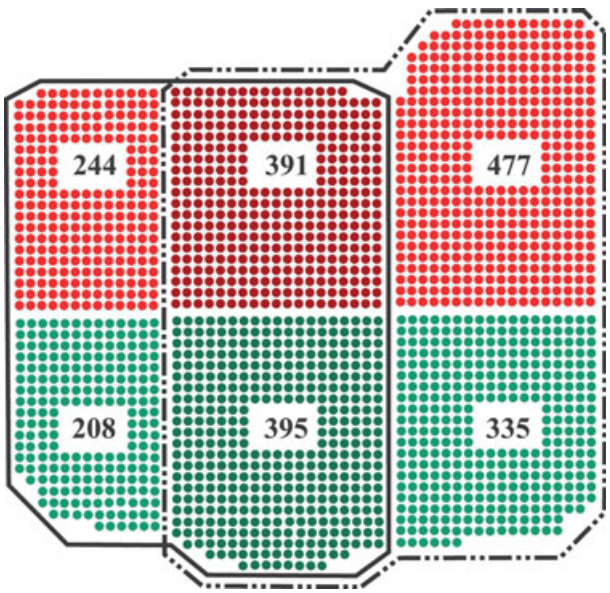

Figure 3 Comparison of differential genes expression in terms of transcript fold changes between microarray and RNA-seq data derived from hydrogen-producing *Chlamydomonas reinhardtii* cells. Plotted are mean values (log2 ratios) over all time points.

genes encoding the acetylglutamate kinase (*Agk*1; *Cre01.g015000*) and aspartate aminotransferase (*Ast*5) were induced, both initial key enzymes for recycling of amino acids. D-cysteine desulphhydrase (*Acd*1), an enzyme that produces sulphides from cysteine, was also highly abundant, thus providing a cell internal source of sulphur. Another enzyme responsible for redistributing intracellular sulphur by using methionine sulphoxide as substrate is the methionine sulphoxide reductase (*Msrb*2). The corresponding transcript was also highly up-regulated in the late phase of the experiment.

From these data, we could confirm an earlier observed shift within the cellular amino acid composition, with a decreased amount of cysteine and increased alanine concentration during hydrogen production (Doebbe *et al.*, 2010). The amino acids composition can be also affected by N-acetyl-l-glutamate synthase activity (e.g. *Nags*2; *Cre02.g144800*), which is known to promote stress tolerance (Kalamaki *et al.*, 2009). In our experiments, this gene was shown to be down-regulated, which could have increased sensitivity to environmental changes. Consequently, this gene would be a suitable target for genetic engineering in *C. reinhardtii* to construct more robust phototrophic strains.

Gene expression of lipid- and carbohydrate-related gene models

Another response to stress induced by sulphur deprivation and anaerobiosis is the accumulation of storage compounds like starch and lipid (Doebbe *et al.*, 2010; Matthew *et al.*, 2009). Lipid-related genes were up-regulated in regard to biosynthesis; for example, acetyl-coA synthetase (*Acs; Cre23.g765700*) and several unspecific acyltransferases (*Cre05.g248150* and *Cre09.g392300*) demonstrated an increased transcript abundance (Table 1). We found a strong increase in transcripts involved in trigalactosyldiacylglycerol transport (*Tgd*1 and *Tgd*2) expression level, both enzymes involved in lipid metabolism, with a proposed transport function (Li *et al.*, 2012; Lu *et al.*, 2007; Roston *et al.*, 2012; Wang *et al.*, 2012; Xu *et al.*, 2010). Our results further confirmed previous data (Doebbe *et al.*, 2010) that showed that storage of lipids and starch occurs during the shift from aerobiosis to anaerobiosis. We could not confirm the up-regulation of nitrogen-induced lipid-related proteins determined by Boyle *et al.* (2012), such as *Pdat*1 (*Cre02.g106400*), *Dgat*1 (*Cre01.g045900*) and *Dgtt*1 (*Cre12.g557750*). However, we

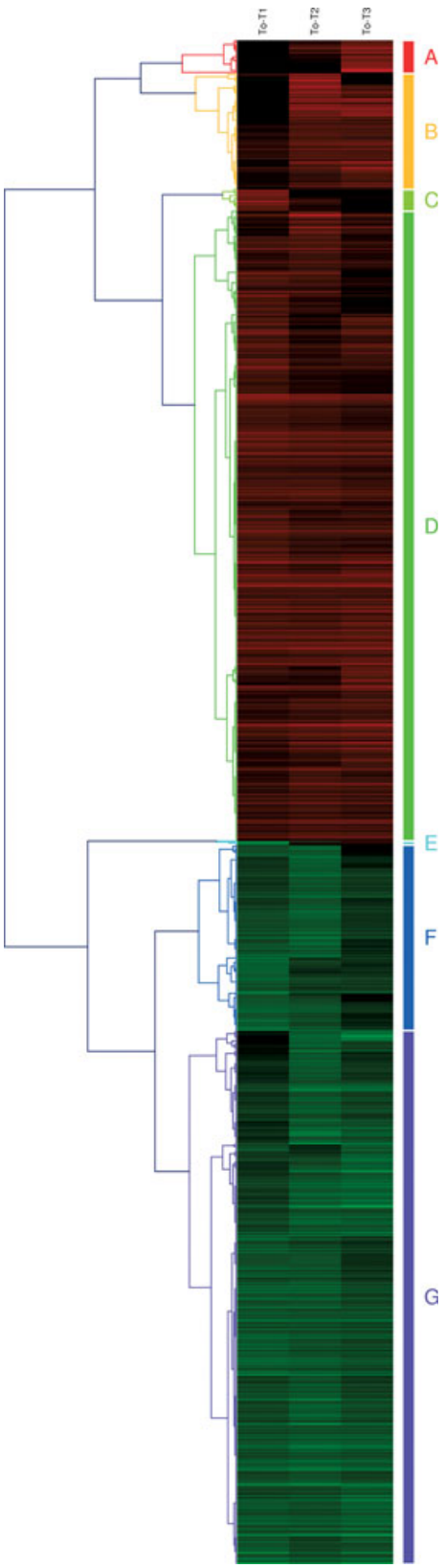

Figure 4 Time-resolved cluster analysis of microarray-based transcript expression level (log2 ratio) for *Chlamydomonas reinhardtii* during hydrogen production (clusters were generated with GENESIS software, Pearson un-centred distance). Red represents up-regulated genes and green down-regulated genes. Gene models, log2 ratios and corresponding clusters are summarized in the Supporting information.

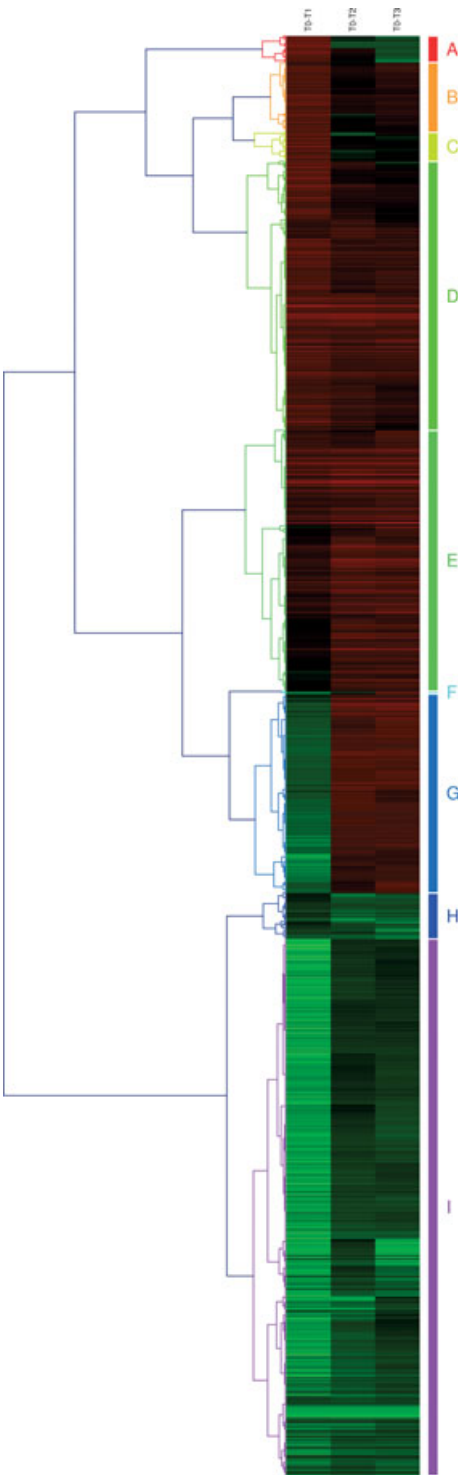

Figure 5 Time-resolved cluster analysis of RNA-seq-based transcript expression level (log2 ratio) for *Chlamydomonas reinhardtii* during hydrogen production (cluster were generated with GENESIS software, Pearson's un-centred distance). Red represents up-regulated genes and green down-regulated genes. Gene models, log2 ratios and corresponding clusters are summarized in Supporting information.

found that *Dgtt2* (*Cre02.g121200*) is induced during the time course of our experiment; therefore, a nutrient-specific induction of *Dgtt* genes can be assumed. Genes involved in biosynthesis of fatty acids were not identified as being differentially regulated

during hydrogen production (e.g. *Kas*1; *Cre22.g765250*, *Kas*2; *Cre07.g335300*). However, fatty acid desaturases (*Cre17.g711150*, *Cre01.g037700*, *Cre16.g672900*, *Cre13.g590500*), creating double carbon bonds, were strongly up-regulated during hydrogen production phase. Accumulation of unsaturated fatty acids was found in *Chlamydomonas* under nitrogen starvation and sulphur starvation (La Russa *et al.*, 2012; Msanne *et al.*, 2012). Further genes related to lipid metabolism were up-regulated in the late phase of the experiment, for example, lipases (*Cre10.g422850*, *Cre07.g322900*).

In regard to starch metabolism, we detected two starch synthase genes with increased expression level (*Cre03.g185250*, *Cre17.g721500*). Genes responsible for starch degradation like isoamylase 3 (*Cre19.g755050*), two alpha amylases (*Cre08.g385500*; *Cre08.g362450*), two phosphorylases (*Cre07.g336950, Cre12.g552200*) and starch debranching (*Cre11.g476650*) also showed an increased expression level during hydrogen production. Additionally, our data demonstrate the induction of glycolysis-related genes during hydrogen-producing conditions. Almost all enzymes of this pathway could be identified (see Table 1 and Figure 6). Most notably, the maximum in average gene expression was observed during peak hydrogen production (Figure 6).

One of the major issues for the survival of the cell under (anaerobic) hydrogen-producing conditions is the maintenance of a balanced ATP/NADP ratio. Genes such as NAD+ kinases (*Nadk*2) (Takahara *et al.*, 2010) were observed to be expressed particularly under these conditions, which are capable of increasing NADP supply. The imbalanced ATP ratio could be reduced by activity of apyrases (e.g. *Cre06.g273500*) that were also induced in our experiments. A major issue for the cell is the recycling of the NADPH, a challenge targeted by the pentose phosphate pathway as well as by the hydrogenase activity and the fermentation processes. All genes of the pentose phosphate pathway and a few fermentative-related genes like alcohol dehydrogenase (*Adh*1; *Cre20.g758200*), pyruvate formate lyase (*Pfl*1; *Cre01.g044800*, just RNA-seq) and acetate kinase (*Ack*1; *Cre09.g396700*) showed increased gene expression (see Mus *et al.* (2007)). These enzymes are the main source for reduction equivalents for the cell under anaerobic stress conditions. Other enzymes that could also provide reduction equivalents are malic

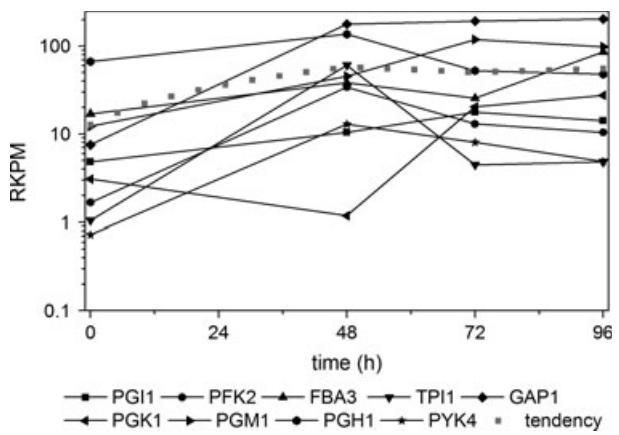

Figure 6 RKPM values calculated from RNA-seq data for all glycolysis-related genes during hydrogen production in *Chlamydomonas reinhardtii*. The dotted grey line determines the tendency of gene expression of all transcript models over time.

enzymes (e.g. *Cre06.g268750*, *Cre14.g629750*, *Cre14.g629700*, *Cre06.g251400*) and isocitrate dehydrogenases (e.g. *Idh*3; *Cre04.g214500*), with all of these gene models showing an increased gene expression during the course of hydrogen production.

Differential gene expression of stress-related gene models

One of the major findings in our data was the strong up-regulation of genes responding to oxidative stress and detoxification (Table 1). Up-regulation of reactive oxygen species (ROS) scavengers such as L-ascorbate peroxidases (*Apx*3 and *Apx*5) and the thylacoidal peroxidase (*SapX*) was detected with RNA-seq and microarrays (Table 1). Such enzymes could be an indication of an induced glutathione–ascorbate cycle, and several studies investigated this cycle (Dietz, 2003, 2010; Maruta *et al.*, 2010; Nagy *et al.*, 2012; Sano *et al.*, 2001; Shigeoka *et al.*, 2002; Steenvoorden and van Henegouwen, 1997; Takeda *et al.*, 1997, 2000; Urzica *et al.*, 2012). However, the mono-dehydroascorbate reductases were not up-regulated (*Mdar*1 and *Mdar*4). Furthermore, the last enzyme in the cycle, the glutathione reductase, was not differentially expressed in our experiments, in contrast to several dehydroascorbate reductases (*Dhar*2 and *Dhar*3); glutathione peroxidases and transferases are also known to be involved in the detoxification process of ROS. ROS stress response regulation is predicted for *Sor*1 (*Cre07.g321550*), a DNA-binding protein, which is up-regulated in our experiment and found by Fischer *et al.* (2012). Detoxification of ROS is predicted for flavodoxin quinone reductases (*Fqr*1-3), which were also up-regulated during this experiment. Additionally, we found numerous genes up-regulated with a potential NADP or FAD oxidoreductase activity (e.g. *Cre07.g346800*, *Cre16.g671450*, *Cre02.g139200*), also known to be potential ROS detoxicants. The oxidative stress is also targeted by the terminal oxygenases (*Pto*1 and *Pto*2), which are involved in carotenoid synthesis (Carol and Kuntz, 2001). Additionally, the alternative oxidase *Aox*1, which is known to be induced under ROS stress in higher plants and involved in nitrogen metabolism in *Chlamydomonas* (Baurain *et al.*, 2003), was observed to be up-regulated. The function for this enzyme during hydrogen production is not known, however, is most likely stress related. In the high hydrogen production mutant *Stm*6 (Kruse *et al.*, 2005), which is highly light sensitive and also partly defective in efficient NPQ (Nguyen *et al.*, 2011), a sharp increase in *Aox* expression levels was also observed. Increased transcript levels of other genes involved in detoxification such as catalase 2 (*Cre09.g417150*) and two superoxide dismutases (*Cre10.g436050*, *Cre02.g096150*) genes were also identified.

Differential gene expression of unknown gene models during peak hydrogen production

Of special interest was the group of unknown differentially expressed genes during hydrogen production, which could be successfully determined within both methods (Table 1). Examples for highly up-regulated genes found in both data sets with putative but not fully defined function and with a high potential as targets for detailed investigations are A glutathione transferase (*Cre12.g559800*), an isoflavone reductase like protein (*Cre11.g477200*), two oxidoreductases (*Cre07.g355500*, *Cre12.g493500*) and several thioredoxins (*Cre01.g057750*, *Cre05.g248500*, *Cre07.g315100*, *Cre14.g624150*). These enzymes are examples for redox stress-induced genes and

therefore potential targets to improve oxidative stress tolerance in *C. reinhardtii*. Furthermore, a potential lipase (*Cre02.g127300*) and a fatty acid desaturase (*Cre17.g711150*) could be identified as being highly up-regulated, which could be of interest in regard to the lipid metabolism in *C. reinhardtii*.

Gene models with potential transport function showed a strong increase in transcript abundance (see Table 1), for example, *Cre16.g656150*, *Cre10.g445000*, *Cre01.g061650* and *Cre03.g166050*. Furthermore, gene models related to phosphate uptake and few carbonic anhydrases [*Cah*1; *Cre09.g405750*, *Cah*7; *Cre13.g607350*, *Cah*9; *Cre05.g243300* and *Cag*1; *Cre12.g516450* (mitochondrial)] also showed an increased transcript level; however, the function of *Cah*7 and *Cah*9 during hydrogen production is unclear, because gene expression could be hardly detected so far (Moroney *et al.*, 2011).

Examples for strong down-regulated genes that could be of interest are mitochondrial carrier proteins (*Cre02.g094250*, *Cre28.g776600*), one NAD-dependent glycerol-3-phosphate dehydrogenase (*Cre01.g053000*) and one acyl carrier protein (*Cre14.g621650*). Differential gene expression of such enzymes indicates potential problems in metabolite transport and communication between compartments in *C. reinhardtii*. We found several other mitochondrial genes, like succinate dehydrogenase or NADH dehydrogenase, which were not significantly affected in our experiment, while gene models encoding subunits of the cytochrome oxidase were strongly up-regulated at all time points. Interestingly, *Cre09.g388150* encoding a mitochondrial translation factor (*Mrpl*36) was highly up-regulated at the late time points. Normally, this factor is involved in translation of mitochondrial proteins (Piao *et al.*, 2009; Prestele *et al.*, 2009).

Finally, we compared gene expression for unknown genes under hydrogen production and sulphur stress (Gonzalez-Ballester *et al.*, 2010; Toepel *et al.*, 2011) and determined several genes that could be identified with numerous reads in all data sets (see Table 1; sulphur-related genes).

Differential gene expression of genes coding for regulatory elements

A time-resolved analysis of RNA-seq data (RPKM values for all differential expressed gene models, plotted in Figure 7) lead to the hypothesis that *C. reinhardtii* first responds to sulphur stress by induction of transport systems, as described earlier (Gonzalez-Ballester *et al.*, 2011; Toepel *et al.*, 2011). However, before the cells reorganize its metabolism, gene expression of specific regulator genes changes. Our data support this hypothesis by demonstrating an early down-regulation of genes (Figure 7a) that are involved in functional assembly of multiprotein complexes, for example the mRNA maturation factor *Nac*2 (*Mbd*2; *Cre02.g124700*). *Nac*2 stabilizes the RNA of the photosystem II subunit *Psb*D and is a key element for the synthesis of photosystem II (Ossenbuhl and Nickelsen, 2000). Other examples for induction of transcription control factors are *Tda*1, which promotes the *Atp*A translation (*Cre08.g358350*, down-regulated), *Mca*1 promoting *Pet*A maturation (*Cre08.g358250*, no change in gene expression), *Tca*1 promoting *Pet*A translation (*Cre09.g415500*, down-regulated), *Mrl*1 promoting *Rbc*L translation (*Cre06.g298300*, down-regulated) and *Mbb*1 promoting the maturation of *Psb*B (*Cre09.g416200*, no change in gene expression). These results clearly demonstrate that control of translation is a first response to stress in *C. reinhardtii* (Eberhard *et al.*, 2011). Proteins involved in regulation of the RUBISCO (*Rmt*1) and potential methyltransferases like *Cre27.g774300*, *Cre08.g368700* and *Cre12.g503800* are also enzymes that regulate protein activity and showed a differential gene expression before complete pathways are remodelled. Flagellar-related genes and genes promoting cell cycle are down-regulated (see Supporting information), indicating a typical stress response and a switch from homoeostasis and cell maintenance to cell survival.

Induction of cell death-related gene models

The hydrogen production phase is characterized by gene expression for genes responsible for stress resistance and NADPH recycling. The final step is the activation of genes inducing apoptosis and cell death. Late up-regulated genes (Figure 7d) are genes related to a specific group of genes and to the cessation of numerous intracellular processes. The end of hydrogen production in *C. reinhardtii* was reached after 96-h cultivation under sulphur deprivation. At this time, the entire cell system starts to degrade. Cell apoptosis could be initiated by differential expression of proteins like DAD1 (Defender against death, *Cre02.g108400*) or APAF1 (apoptotic protease-activating factor 1, found as protein but not annotated in *Chlamydomonas*) (Moharikar *et al.*, 2007).

Activating proteins involved in protein degradation like ubiquitins (*Ubq*1, *Cre18.g750000*, *Ubq*7; *Cre13.g563600*), ubiquitin protein ligases and ubiquitin protein proteases (at least five genes are up-regulated late during the experiments, e.g. *Cre06.g266350*, *Cre12.g533750*, *Cre08.g364550* (Figure 7d)); furthermore, autophagy-inducing genes were also up-regulated like *Apg*8 (*Cre16.g689650*) (Perez-Perez *et al.*, 2010). Finally, late induced proteases (*Cre16.g663350*, *Clpp*2; *Cre12.g521450*, *Cre06.g274700*, *Deg*11; *Cre12.g498500*, *Deg*1; *Cre02.g088400*, *Cep*1; *Cre09.g407700*, *Cep*2; *Cre05.g247800*) lead to cell death (see Figure 8). The up-regulation of such enzymes as final step under sulphur starvation was also described earlier (Gonzalez-Ballester *et al.*, 2010).

Discussion

Hydrogen production in *C. reinhardtii* has been a major target of research, and during the last years, intensive systematic analysis of the hydrogen production metabolism resulted in a remarkable increase in knowledge regarding the process in general (Eroglu and Melis, 2011; Kruse *et al.*, 2005; Nguyen *et al.*, 2011). As a consequence, the understanding how cells remodel their metabolism and their ability to compensate mutations of specific genes (knock out, RNAi, etc.) dramatically increased. Strong examples are the role of fermentative-related genes (Catalanotti *et al.*, 2012; Grossman *et al.*, 2011; Magneschi *et al.*, 2012) and neutral lipids during sulphur deprivation (La Russa *et al.*, 2012). To date, hydrogen production rates are still below viable relevance to the alternative fuel production and must be improved. The H_2 production phase is still too short, and prolongation of the process is required to increase system yields. A technical solution in this regard was recently provided (Lehr *et al.*, 2012) by inducing hydrogen production phase by controlled and limited microsupply of sulphur. The data presented here demonstrate the possibility to use transcriptome analysis techniques to determine gene targets whose expression levels are crucial for efficient hydrogen production. Our results clearly show that both methods, RNA-seq and microarray, were able to provide complimentary data to further elucidate regulation of genes in *C. reinhardtii* during hydrogen production and completed our picture regarding differential gene expression for

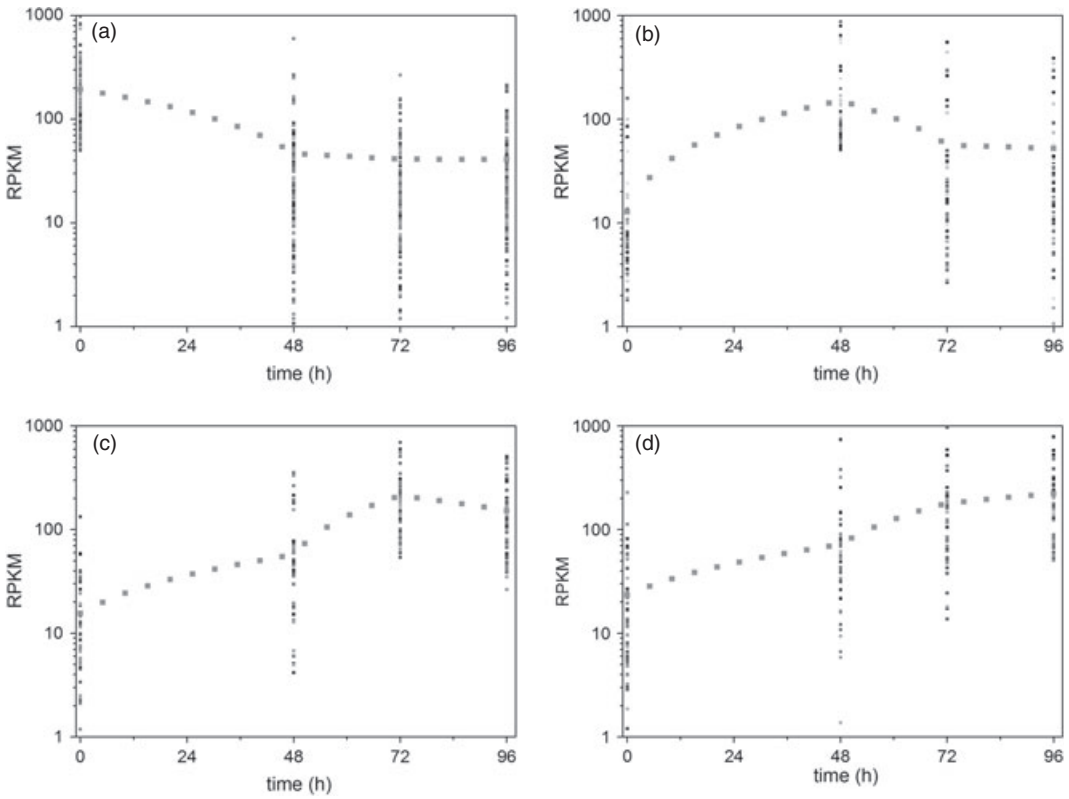

Figure 7 Grouping of gene models according to the time point of highest transcript level during hydrogen production in *Chlamydomonas reinhardtii*. Plotted are the RKPM values (determined with RNA-seq) of all differential expressed gene models at each time point (a: maximum after 0 h, b: maximum after 48 h, c: maximum after 72 h and d: maximum after 96 h); the dotted grey lines represent the tendency of all gene expression of models during the time course.

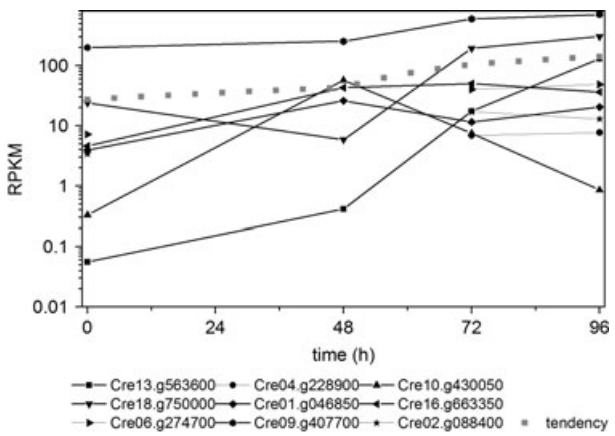

Figure 8 RKPM values calculated from RNA-seq data of several ubiquitins and proteases during hydrogen production in *Chlamydomonas reinhardtii*. The dotted grey line determines the tendency of gene expression of all transcript models over time.

adaptation to sulphur stress, anaerobiosis and induction of semi-constant hydrogen production. The sulphur deprivation–related results are in very good accordance with previously published data (Gonzalez-Ballester *et al.*, 2010; Toepel *et al.*, 2011). Indeed, the expression of sulphur-related transport systems, redistribution of internal sulphur sources and gene expression responsible for accumulation of storage, for example, lipids and starch compounds, were verified with our experiments. We could additionally identify gene expression for distinct genes in regard to oxidative stress, photosynthesis and general metabolism. Furthermore, the larger dynamic range of the RNA-seq technique, as earlier already postulated by Gonzalez-Ballester *et al.* (2010), was confirmed here. It is noteworthy to mention that RNA-seq enabled for an improved time course-dependent analysis of gene expression, while the microarray data could only be used to identify relative transcript abundance variation that remained for most of the genes during course of the experiment.

The results furthermore clearly demonstrate that hydrogen production forces *C. reinhardtii* cells to deal with two major stressors, ROS and maintenance of intracellular energy balances. This is reflected by expression of genes responsible for transport systems and for remodelling metabolic pathways necessary to access internal sulphur sources. ROS accelerate the process of cell destruction and must be efficiently removed from the system, a strategy that is probably targeted by over-expression of genes coding protection-related proteins. The consequence of a better control of cellular answers to ROS is the maintenance of a minimum of photosystem II activity and photosynthetic electron transport, so that electron supply to the hydrogenase can be prolonged. Consequently, genes for targeted genetic manipulation are members of the glutathione–ascorbate cycle (such as ascorbate peroxidases) and genes coding for superoxide dismutases were shown to be up-regulated in the late phase of hydrogen production (Figure 8).

A prolonged hydrogen production as a result of partial PS II protection, for example, by efficient xanthophyll cycle, was already described by others (Scoma *et al.*, 2012; Torzillo *et al.*,

2009). Furthermore, the strong down-regulation of the whole cytochrome b_6f complex implicates a potential bottleneck in electron transport in this photosynthetic multiprotein subunit. A stabilization of this complex could therefore also be a target of future studies. In this regard, an increased consumption of reduction equivalents by such enzymes could be problematic; however, a stable electron transport rate is crucial for biotechnological applications targeting to improve hydrogen production in *C. reinhardtii*. One of the main goals in future projects should be the generation of *C. reinhardtii* strains with stable Cyt-b_6f- and PS I systems under stress conditions, even with low electron transport capacities. Philipps *et al.* (2012) demonstrated in this regard the effect of limited electron flow towards the hydrogenase in nitrogen-deprived *C. reinhardtii* cultures, with the consequence of low hydrogen production rates.

In addition, cell internal energy distribution has to be directed towards the hydrogenase, which can be realized by promoting efficient storage breakdown, minimizing competition reactions and providing reduction equivalents to the enzyme. Providing hydrogenase with electrons and protons is essential for productivity; therefore, NADPH and/or ferredoxin recycling is crucial (Winkler *et al.*, 2010). Our data indicate that the cells most likely try to balance the ATP/NADPH ratio by induction of several alternative pathways and enzymes. However, a reduced activity of transport mechanisms and a disturbed interorganelle communication between chloroplast and mitochondria most likely reduce the functionality of the cellular system during the hydrogen production process. In addition, an increased competition between fermentation processes and hydrogenases for electrons reduces the capacity for hydrogen production.

Lipid degradation is strongly reduced under anaerobic conditions; therefore, accumulation of neutral lipids has to be avoided to increase H_2 production rates as the lipid storage sink will be hardly available as a substrate for hydrogenase (see also (Doebbe *et al.*, 2010; Miller *et al.*, 2010; Philipps *et al.*, 2012)). It has been also previously shown how accumulation of starch and lipids influences hydrogen production capacities in *C. reinhardtii* under different nutrient starvation conditions (Philipps *et al.*, 2012). Within this work, several genes were identified to be involved in accumulation/degradation of polar lipids, being also a potential target for manipulation, with the goal to reduce energy storage as lipids. Difficulties of direct manipulation of potential key enzymes of the lipid metabolism were however demonstrated previously (La Russa *et al.*, 2012), thus demonstrating metabolomic complexity and remodelling capacity of a cellular metabolism; therefore, construction of a complete metabolomic network is essential for a precise prediction of gene manipulation effects. One example for metabolomic complexity is the group of malic enzymes, proteins which are not directly related to lipid chemistry, but are known to affect lipid accumulation. Over-expression of malic enzymes increased lipid concentration in fungi; however, at the same time, these enzymes produce reduction equivalents, useful for hydrogen production in *C. reinhardtii* (Zhang *et al.*, 2007). Malic enzymes, the pentose phosphate pathway and isocitrate dehydrogenases are main sources for metabolomic reduction equivalents during hydrogen production, and stabilization of function for these enzymes could also extend hydrogen production. However, prior to such engineering approaches, precise full energy balance analyses of potential future biorefineries are needed, because a high lipid-containing biomass (*C. reinhardtii* after hydrogen production) could also be useful in follow-up processes in which the remaining biomass is used to produce biomethane via fermentation (Mussgnug *et al.*, 2010). The rate of gene expression of distinct lipases, as shown here (Table 1), is not sufficient for lipid degradation during hydrogen production, because the lack of oxygen prevents beta oxidation. Crucial for storage compound breakdown is also the activity of the respiratory chain in the mitochondria. Our data indicate that the transport of metabolites between the compartments is strongly affected during the time course of the experiments. The importance of a functioning proton pumping in *C. reinhardtii* during hydrogen production has been already described (Lecler *et al.*, 2011). Therefore, it is feasible to suggest that stabilization of metabolite transport and enhanced degradation of starch and a reduced lipid accumulation should increase the hydrogen production phase significantly.

In addition, this study identifies a number of unknown genes that are highly up- or down-regulated during hydrogen production. The function of most of these genes can only be assumed, and characterization needs to be carried out in future studies to reveal the possible function during hydrogen production.

Our data furthermore demonstrated that mutation of *lhcsr*3 in *Npq*4 did only slightly improve hydrogen production in *Chlamydomonas* most likely due to the fact that stress-induced damage of photosynthesis during of sulphur deprivation phase was not increased compared to the corresponding parental strain. However, as it is known that NPQ is an important factor in regard to hydrogen production (Kruse *et al.*, 2005; Nguyen *et al.*, 2011), further detailed analysis of the different LHCSR isoforms and especially LHCSR1 (gene expression in *npq*4 was identical to wild type) is necessary to unravel the function of these enzymes during hydrogen production.

The comparison of results derived from both methods, microarray and RNA-seq, showed a good accordance. The higher dynamic range and capacity to detect unknown genes however clearly promotes the usage of transcript sequencing. Nevertheless, there are some critical factors in regard of RNA-seq data: The calculation of fold changes of stress-induced genes, where no reads could be determined in the control sample, is problematic. Furthermore, one has to determine whether the gene of interest was not expressed in the control sample or whether the sequences were not amplified. Furthermore, RNA-seq still could underestimate low abundant genes, for example, fermentative pathways genes. And finally, alignment and differentiation of reads for transcripts with high similarity, for example, LHCBM proteins or histones, is difficult and has to be analysed for each gene in detail by RT-Q-PCR. For the LHCBM proteins, we could just use a small amount of sequences specific to each isoform (for detailed explanations see Fang *et al.* (2012)). On the other hand, microarrays provide an easy-to-use platform with high reproducibility and analysis can be performed in a standardized way. Therefore, a combined application of both methods could be considered as a powerful strategy to achieve profound and deep insights into distinct cellular metabolic decisions.

Experimental procedure

Cultures, growth and H_2 production conditions

As reference strains, we used the wild-type 4A+ and the non-photochemical quenching mutant *npq*4, which was generated by

insertional mutagenesis, resulting in a knockout of the *LHCSR3.1* and *LHCSR3.2* genes (Niyogi *et al.*, 1997; Peers *et al.*, 2009). Cultures were grown mixotrophically in standard TAP media (Harris, 1989) until early stationary phase in constant light (50 μE/m^2/s) at 25–30 °C, harvested via centrifugation and resuspended after washing (three times) in TAP minus S media (Melis and Happe, 2001). For H_2 production measurements, cells were cultivated for 96 h in constant light (200 μE/m^2/s^1) in sealed 300 mL bioreactors. Hydrogen production was measured as volume determination of gas production, and gas quality was measured by gas chromatography as earlier described (Doebbe *et al.*, 2010). Photosynthetic quantum yields were monitored with a MINI-PAM (Walz, Germany) (Nguyen *et al.*, 2011).

Sample collection, RNA preparation and transcriptomics

Samples of *C. reinhardtii* from bioreactor-cultivated cultures (300 mL) were collected at 24, 48, 72 and 96 h after sulphur deprivation (T1, T2, T3 and T4). We excluded the time point T1 (24 h) from analysis and comparison, because a high amplification of GC-rich reads in RNA-seq was detected and adapted the labelling: 48 h (T1), 72 h (T2) and 96 h (T3) for both methods. Reference samples (T0) were harvested from cultures of the corresponding strain prior to sulphur deprivation. Samples were immediately centrifuged at 8300 ***g*** for 2 min at room temperature, cell pellets were immediately lysed with RNA lysis buffer and RNA was isolated as previously described (Nguyen *et al.*, 2011). RNA samples were pooled from three independent experiments and used for microarray (three technical replicates) and RNA-seq analyses. RNA quality was determined with an Agilent© Bioanalyzer (Agilent Inc., Santa Clara, CA). *Chlamydomonas reinhardtii* microarray slides (Agilent© 4 × 44 k, no: 024664) were used for the transcript analyses (see Toepel *et al.* (2011) for details). RNA labelling (Quick RNA amplification and labelling kit; Agilent©) and microarray hybridization (16 h at 60 °C) were carried out according to the supplied manual. Microarrays were washed after hybridization according to the Agilent© manual, dried in a centrifuge and scanned with a 5-μm resolution in an Agilent© DNA microarray scanner. Data extraction was achieved using the feature extraction software (10.7.3.1; Agilent©), and data were normalized and analysed using the software EMMA2 (see Toepel *et al.* (2011)). We used a robust normalization method (lowess), performed significance tests within all experiments and considered only those probes showing a significant change in their expression (*P*-values smaller than 0.05). To further limit our result set, we included only genes that demonstrated at least a twofold up- or down-regulation. RNA-seq was performed as described by (Illumina Inc., San Diego, CA). The cDNA libraries were assembled according to Illumina's RNA-seq protocol, loaded and sequenced as 36-mers as single reads. Raw sequence files were aligned against *Chlamydomonas* genome (phytozome 4.3) with a new developed program, SARUMAN (Blom *et al.*, 2011), which allows rapid and precise alignment. We allowed a 6% error in the alignments and analysed the best position for each read. Without losing all other positions, only unique reads per gene model were used for the determination of the expression of a gene. Expression estimates were obtained for each individual run in units of RPKMs (reads per kilobase of mappable transcript length per million mapped reads) after normalization by the number of aligned reads and map-able transcripts (Boyle *et al.*, 2012). Based on the normalized RPKM values, we estimated *log*2 ratios for the gene models. To calculate newly induced genes, characterized by zero reads in the control samples, we used average minimal expression level as control value. All transcriptomic data were clustered and visualized with the GENESIS software (Sturn *et al.*, 2002) using Pearson's un-centred distance. Significance test was performed with EdgeR software (http://bioconductor.org/packages/2.10/bioc/html/edgeR.html).

Validation of RNA-seq and microarray data was performed for reference genes with quantitative RT-PCR as described (Nguyen *et al.*, 2011).

Microarray data can be accessed at NCBI (Geo): GSE41728.

RNA-seq data can be accessed at ArrayExpress: E-MTAB-1329.

Acknowledgements

The authors thank the Federal Ministry of Science [BMBF ForSys Partner grant contract 0315265A] and the EU/Energy FP7 project SOLAR-H_2 [contract 212508], for financial support and Prof. Krishna Niyogi from the University of California for providing the *npq4* and *4A+* mutants. Patrick May was funded by the German Federal Ministry of Education and Research, Systems Biology Research Initiative "GoFORSYS". We thank Christian Rückert (Bielefeld University) for the help with the RNA-seq experiments.

References

Baurain, D., Dinant, M., Coosemans, N. and Matagne, R.F. (2003) Regulation of the alternative oxidase Aox1 gene in *Chlamydomonas reinhardtii*. Role of the nitrogen source on the expression of a reporter gene under the control of the Aox1 promoter. *Plant Physiol.* **131**, 1418–1430.

Blom, J., Jakobi, T., Doppmeier, D., Jaenicke, S., Kalinowski, J., Stoye, J. and Goesmann, A. (2011) Exact and complete short-read alignment to microbial genomes using Graphics Processing Unit programming. *Bioinformatics*, **27**, 1351–1358.

Bonente, G., Ballottari, M., Truong, T.B., Morosinotto, T., Ahn, T.K., Fleming, G.R., Niyogi, K.K. and Bassi, R. (2011) Analysis of LhcSR3, a protein essential for feedback de-excitation in the green alga *Chlamydomonas reinhardtii*. *PLoS Biol.* **9**, 107–110.

Boyle, N.R., Page, M.D., Liu, B., Blaby, I.K., Casero, D., Kropat, J., Cokus, S.J., Hong-Hermesdorf, A., Shaw, J., Karpowicz, S.J., Gallaher, S.D., Johnson, S., Benning, C., Pellegrini, M., Grossman, A. and Merchant, S.S. (2012) Three acyltransferases and nitrogen-responsive regulator are implicated in nitrogen starvation-induced triacylglycerol accumulation in *Chlamydomonas*. *J. Biol. Chem.* **287**, 15811–15825.

Carol, P. and Kuntz, M. (2001) A plastid terminal oxidase comes to light: implications for carotenoid biosynthesis and chlororespiration. *Trends Plant Sci.* **6**, 31–36.

Catalanotti, C., Dubini, A., Subramanian, V., Yang, W.Q., Magneschi, L., Mus, F., Seibert, M., Posewitz, M.C. and Grossman, A.R. (2012) Altered fermentative metabolism in *Chlamydomonas reinhardtii* mutants lacking pyruvate formate lyase and both pyruvate formate lyase and alcohol dehydrogenase. *Plant Cell*, **24**, 692–707.

Chen, M., Zhao, L., Sun, Y.L., Cui, S.X., Zhang, L.F., Yang, B., Wang, J., Kuang, T.Y. and Huang, F. (2010) Proteomic analysis of hydrogen photoproduction in sulfur-deprived *Chlamydomonas* cells. *J. Proteome Res.* **9**, 3854–3866.

Chochois, V., Dauvillee, D., Beyly, A., Tolleter, D., Cuine, S., Timpano, H., Ball, S., Cournac, L. and Peltier, G. (2009) Hydrogen production in *Chlamydomonas*: photosystem II-dependent and -independent pathways differ in their requirement for starch metabolism. *Plant Physiol.* **151**, 631–640.

Chochois, V., Constans, L., Dauvillee, D., Beyly, A., Soliveres, M., Ball, S., Peltier, G. and Cournac, L. (2010) Relationships between PSII-independent hydrogen bioproduction and starch metabolism as evidenced from isolation of starch catabolism mutants in the green alga *Chlamydomonas reinhardtii*. *Int. J. Hydrogen Energy*, **35**, 10731–10740.

Desplats, C., Mus, F., Cuine, S., Billon, E., Cournac, L. and Peltier, G. (2009) Characterization of *Nda2*, a plastoquinone-reducing type II NAD(P)H dehydrogenase in *Chlamydomonas* chloroplasts. *J. Biol. Chem.* **284**, 4148–4157.

Dietz, K.J. (2003) Redox control, redox signaling, and redox homeostasis in plant cells. *Int. Rev. Cytol.* **228**, 141–193.

Dietz, K.J. (2010) Redox-dependent regulation, redox control and oxidative damage in plant cells subjected to abiotic stress. *Methods Mol. Biol.* **639**, 57–70.

Doebbe, A., Keck, M., La Russa, M., Mussgnug, J.H., Hankamer, B., Tekce, E., Niehaus, K. and Kruse, O. (2010) The interplay of proton, electron, and metabolite supply for photosynthetic H_2 production in *Chlamydomonas reinhardtii*. *J. Biol. Chem.* **285**, 30247–30260.

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.

Eberhard, S., Loiselay, C., Drapier, D., Bujaldon, S., Girard-Bascou, J., Kuras, R., Choquet, Y. and Wollman, F.A. (2011) Dual functions of the nucleus-encoded factor TDA1 in trapping and translation activation of *atp*A transcripts in *Chlamydomonas reinhardtii* chloroplasts. *Plant J.* **67**, 1055–1066.

Eroglu, E. and Melis, A. (2011) Photobiological hydrogen production: recent advances and state of the art. *Bioresour. Technol.* **102**, 8403–8413.

Esper, B., Badura, A. and Rogner, M. (2006) Photosynthesis as a power supply for (bio-)hydrogen production. *Trends Plant Sci.* **11**, 543–549.

Fang, W., Si, Y., Douglass, S., Casero, D., Merchant, S.S., Pellegrini, M., Ladunga, I., Liu, P. and Spalding, M.H. (2012) Transcriptome-wide changes in *Chlamydomonas reinhardtii* gene expression regulated by carbon dioxide and the CO_2-concentrating mechanism regulator CIA5/CCM1. *Plant Cell*, **24**, 1876–1893.

Fischer, B.B., Ledford, H.K., Wakao, S., Huang, S.G., Casero, D., Pellegrini, M., Merchant, S.S., Koller, A., Eggen, R.I.L. and Niyogi, K.K. (2012) SINGLET OXYGEN RESISTANT 1 links reactive electrophile signaling to singlet oxygen acclimation in *Chlamydomonas reinhardtii*. *Proc. Natl Acad. Sci. USA*, **109**, E1302–E1311.

Gonzalez-Ballester, D., Casero, D., Cokus, S., Pellegrini, M., Merchant, S.S. and Grossman, A.R. (2010) RNA-seq analysis of sulfur-deprived *Chlamydomonas* cells reveals aspects of acclimation critical for cell survival. *Plant Cell*, **22**, 2058–2084.

Gonzalez-Ballester, D., Pootakham, W., Mus, F., Yang, W., Catalanotti, C., Magneschi, L., de Montaigu, A., Higuera, J.J., Prior, M., Galvan, A., Fernandez, E. and Grossman, A.R. (2011) Reverse genetics in Chlamydomonas: a platform for isolating insertional mutants. *Plant Methods*, **7**, 24.

Grossman, A.R., Catalanotti, C., Yang, W.Q., Dubini, A., Magneschi, L., Subramanian, V., Posewitz, M.C. and Seibert, M. (2011) Multiple facets of anoxic metabolism and hydrogen production in the unicellular green alga *Chlamydomonas reinhardtii*. *New Phytol.* **190**, 279–288.

Happe, T. and Kaminski, A. (2002) Differential regulation of the Fe-hydrogenase during anaerobic adaptation in the green alga *Chlamydomonas reinhardtii*. *Eur. J. Biochem.* **269**, 1022–1032.

Happe, T., Hemschemeier, A., Winkler, M. and Kaminski, A. (2002) Hydrogenases in green algae: do they save the algae's life and solve our energy problems? *Trends Plant Sci.* **7**, 246–250.

Harris, E.H. (1989) *The Chlamydomonas Sourcebook: A Comprehensive Guide to Biology and Laboratory Use*. Oxford, UK: Academic Press.

Hemschemeier, A. and Happe, T. (2011) Alternative photosynthetic electron transport pathways during anaerobiosis in the green alga *Chlamydomonas reinhardtii*. *Biochim. Biophys. Acta*, **1807**, 919–926.

Hemschemeier, A., Fouchard, S., Cournac, L., Peltier, G. and Happe, T. (2008a) Hydrogen production by *Chlamydomonas reinhardtii*: an elaborate interplay of electron sources and sinks. *Planta*, **227**, 397–407.

Hemschemeier, A., Jacobs, J. and Happe, T. (2008b) Biochemical and physiological characterization of the pyruvate formate-lyase *Pfl*1 of *Chlamydomonas reinhardtii*, a typically bacterial enzyme in a eukaryotic alga. *Eukaryot. Cell*, **7**, 518–526.

Kalamaki, M.S., Alexandrou, D., Lazari, D., Merkouropoulos, G., Fotopoulos, V., Pateraki, I., Aggelis, A., Carrillo-Lopez, A., Rubio-Cabetas, M.J. and Kanellis, A.K. (2009) Over-expression of a tomato N-acetyl-L-glutamate synthase gene (SlNAGS1) in *Arabidopsis thaliana* results in high ornithine levels and increased tolerance in salt and drought stresses. *J. Exp. Bot.* **60**, 1859–1871.

Kruse, O., Rupprecht, J., Bader, K.P., Thomas-Hall, S., Schenk, P.M., Finazzi, G. and Hankamer, B. (2005) Improved photobiological H_2 production in engineered green algal cells. *J. Biol. Chem.* **280**, 34170–34177.

La Russa, M., Bogen, C., Uhmeyer, A., Doebbe, A., Filippone, E., Kruse, O. and Mussgnug, J.H. (2012) Functional analysis of three type-2 DGAT homologue genes for triacylglycerol production in the green microalga *Chlamydomonas reinhardtii*. *J. Biotechnol.* **162**, 13–20.

Lecler, R., Godaux, D., Vigeolas, H., Hiligsmann, S., Thonart, P., Franck, F., Cardol, P. and Remacle, C. (2011) Functional analysis of hydrogen photoproduction in respiratory-deficient mutants of *Chlamydomonas reinhardtii*. *Int. J. Hydrogen Energy*, **36**, 9562–9570.

Lehr, F., Morweiser, M., Sastre, R.R., Kruse, O. and Posten, C. (2012) Process development for hydrogen production with *Chlamydomonas reinhardtii* based on growth and product formation kinetics. *J. Biotechnol.* **162**, 89–96.

Li, Z., Gao, J., Benning, C. and Sharkey, T.D. (2012) Characterization of photosynthesis in *Arabidopsis* ER-to-plastid lipid trafficking mutants. *Photosynth. Res.* **112**, 49–61.

Liu, S., Lin, L., Jiang, P., Wang, D. and Xing, Y. (2011) A comparison of RNA-Seq and high-density exon array for detecting differential gene expression between closely related species. *Nucleic Acids Res.* **39**, 578–588.

Lu, B., Xu, C., Awai, K., Jones, A.D. and Benning, C. (2007) A small ATPase protein of *Arabidopsis*, TGD3, involved in chloroplast lipid import. *J. Biol. Chem.* **282**, 35945–35953.

Magneschi, L., Catalanotti, C., Subramanian, V., Dubini, A., Yang, W.Q., Mus, F., Posewitz, M.C., Seibert, M., Perata, P. and Grossman, A.R. (2012) A mutant in the ADH1 gene of *Chlamydomonas reinhardtii* elicits metabolic restructuring during anaerobiosis. *Plant Physiol.* **158**, 1293–1305.

Malone, J.H. and Oliver, B. (2011) Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol.* **9**, 34.

Maruta, T., Tanouchi, A., Tamoi, M., Yabuta, Y., Yoshimura, K., Ishikawa, T. and Shigeoka, S. (2010) *Arabidopsis* chloroplastic ascorbate peroxidase isoenzymes play a dual role in photoprotection and gene regulation under photooxidative stress. *Plant Cell Physiol.* **51**, 190–200.

Matthew, T., Zhou, W.X., Rupprecht, J., Lim, L., Thomas-Hall, S.R., Doebbe, A., Kruse, O., Hankamer, B., Marx, U.C., Smith, S.M. and Schenk, P.M. (2009) The metabolome of *Chlamydomonas reinhardtii* following induction of anaerobic H_2 production by sulfur depletion. *J. Biol. Chem.* **284**, 23415–23425.

Melis, A. and Happe, T. (2001) Hydrogen production. Green algae as a source of energy. *Plant Physiol.* **127**, 740–748.

Miller, R., Wu, G., Deshpande, R.R., Vieler, A., Gartner, K., Li, X., Moellering, E.R., Zauner, S., Cornish, A.J., Liu, B., Bullard, B., Sears, B.B., Kuo, M.H., Hegg, E.L., Shachar-Hill, Y., Shiu, S.H. and Benning, C. (2010) Changes in transcript abundance in *Chlamydomonas reinhardtii* following nitrogen deprivation predict diversion of metabolism. *Plant Physiol.* **154**, 1737–1752.

Moharikar, S., D'Souza, J.S. and Rao, B.J. (2007) A homologue of the defender against the apoptotic death gene (*dad*1)in UV-exposed *Chlamydomonas* cells is downregulated with the onset of programmed cell death. *J. Biosci.* **32**, 261–270.

Moroney, J.V., Ma, Y., Frey, W.D., Fusilier, K.A., Pham, T.T., Simms, T.A., DiMario, R.J., Yang, J. and Mukherjee, B. (2011) The carbonic anhydrase isoforms of *Chlamydomonas reinhardtii*: intracellular location, expression, and physiological roles. *Photosynth. Res.* **109**, 133–149.

Msanne, J., Xu, D., Konda, A.R., Casas-Mollano, J.A., Awada, T., Cahoon, E.B. and Cerutti, H. (2012) Metabolic and gene expression changes triggered by nitrogen deprivation in the photoautotrophically grown microalgae *Chlamydomonas reinhardtii* and *Coccomyxa sp*. C-169. *Phytochemistry*, **75**, 50–59.

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.

Mussgnug, J.H., Klassen, V., Schluter, A. and Kruse, O. (2010) Microalgae as substrates for fermentative biogas production in a combined biorefinery concept. *J. Biotechnol.* **150**, 51–56.

Nagy, V., Tengolics, R., Schansker, G., Rakhely, G., Kovacs, K.L., Garab, G. and Toth, S.Z. (2012) Stimulatory effect of ascorbate, the alternative electron donor of photosystem II, on the hydrogen production of sulphur-deprived *Chlamydomonas reinhardtii*. *Int. J. Hydrogen Energy*, **37**, 8864–8871.

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) Transcriptome for photobiological hydrogen production induced by sulfur deprivation in the green alga *Chlamydomonas reinhardtii*. *Eukaryot. Cell*, **7**, 1965–1979.

Nguyen, A.V., Toepel, J., Burgess, S., Uhmeyer, A., Blifernez, O., Doebbe, A., Hankamer, B., Nixon, P., Wobbe, L. and Kruse, O. (2011) Time-course global expression profiles of *Chlamydomonas reinhardtii* during photo-biological H_2 production. *PLoS ONE*, **6**, e29364.

Niyogi, K.K., Bjorkman, O. and Grossman, A.R. (1997) *Chlamydomonas* xanthophyll cycle mutants identified by video imaging of chlorophyll fluorescence quenching. *Plant Cell*, **9**, 1369–1380.

Oshlack, A., Robinson, M.D. and Young, M.D. (2010) From RNA-seq reads to differential expression results. *Genome Biol.* **11**, 220.

Ossenbuhl, F. and Nickelsen, J. (2000) cis- and trans-Acting determinants for translation of *psb*D mRNA in *Chlamydomonas reinhardtii*. *Mol. Cell. Biol.* **20**, 8134–8142.

Peers, G., Truong, T.B., Ostendorf, E., Busch, A., 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. *Nature*, **462**, 518–521.

Perez-Perez, M.E., Florencio, F.J. and Crespo, J.L. (2010) Inhibition of target of rapamycin signaling and stress activate autophagy in *Chlamydomonas reinhardtii*. *Plant Physiol.* **152**, 1874–1888.

Philipps, G., Krawietz, D., Hemschemeier, A. and Happe, T. (2011) A pyruvate formate lyase-deficient *Chlamydomonas reinhardtii* strain provides evidence for a link between fermentation and hydrogen production in green algae. *Plant J.* **66**, 330–340.

Philipps, G., Happe, T. and Hemschemeier, A. (2012) Nitrogen deprivation results in photosynthetic hydrogen production in *Chlamydomonas reinhardtii*. *Planta*, **235**, 729–745.

Piao, L., Li, Y., Kim, S.J., Byun, H.S., Huang, S.M., Hwang, S.K., Yang, K.J., Park, K.A., Won, M., Hong, J., Hur, G.M., Seok, J.H., Shong, M., Cho, M.H., Brazil, D.P., Hemmings, B.A. and Park, J. (2009) Association of LETM1 and MRPL36 contributes to the regulation of mitochondrial ATP production and necrotic cell death. *Cancer Res.* **69**, 3397–3404.

Prestele, M., Vogel, F., Reichert, A.S., Herrmann, J.M. and Ott, M. (2009) Mrpl36 is important for generation of assembly competent proteins during mitochondrial translation. *Mol. Biol. Cell*, **20**, 2615–2625.

Ramskold, D., Kavak, E. and Sandberg, R. (2012) How to analyze gene expression using RNA-sequencing data. *Methods Mol. Biol.* **802**, 259–274.

Roston, R.L., Gao, J., Murcha, M.W., Whelan, J. and Benning, C. (2012) TGD1, -2, and -3 proteins involved in lipid trafficking form ATP-binding cassette (ABC) transporter with multiple substrate-binding proteins. *J. Biol. Chem.* **287**, 21406–21415.

Sano, S., Ueda, M., Kitajima, S., Takeda, T., Shigeoka, S., Kurano, N., Miyachi, S., Miyake, C. and Yokota, A. (2001) Characterization of ascorbate peroxidases from unicellular red alga *Galdieria partita*. *Plant Cell Physiol.* **42**, 433–440.

Scoma, A., Krawietz, D., Faraloni, C., Giannelli, L., Happe, T. and Torzillo, G. (2012) Sustained H_2 production in a *Chlamydomonas reinhardtii* D1 protein mutant. *J. Biotechnol.* **157**, 613–619.

Shigeoka, S., Ishikawa, T., Tamoi, M., Miyagawa, Y., Takeda, T., Yabuta, Y. and Yoshimura, K. (2002) Regulation and function of ascorbate peroxidase isoenzymes. *J. Exp. Bot.* **53**, 1305–1319.

Smith, D.R., Quinlan, A.R., Peckham, H.E., Makowsky, K., Tao, W., Woolf, B., Shen, L., Donahue, W.F., Tusneem, N., Stromberg, M.P., Stewart, D.A., Zhang, L., Ranade, S.S., Warner, J.B., Lee, C.C., Coleman, B.E., Zhang, Z., McLaughlin, S.F., Malek, J.A., Sorenson, J.M., Blanchard, A.P., Chapman, J., Hillman, D., Chen, F., Rokhsar, D.S., McKernan, K.J., Jeffries, T.W., Marth, G.T. and Richardson, P.M. (2008) Rapid whole-genome mutational profiling using next-generation sequencing technologies. *Genome Res.* **18**, 1638–1642.

Steenvoorden, D.P. and van Henegouwen, G.M. (1997) The use of endogenous antioxidants to improve photoprotection. *J. Photochem. Photobiol. B Biol.* **41**, 1–10.

Sturn, A., Quackenbush, J. and Trajanoski, Z. (2002) Genesis: cluster analysis of microarray data. *Bioinformatics*, **18**, 207–208.

Takahara, K., Kasajima, I., Takahashi, H., Hashida, S.N., Itami, T., Onodera, H., Toki, S., Yanagisawa, S., Kawai-Yamada, M. and Uchimiya, H. (2010) Metabolome and photochemical analysis of rice plants overexpressing *Arabidopsis* NAD kinase gene. *Plant Physiol.* **152**, 1863–1873.

Takeda, T., Ishikawa, T. and Shigeoka, S. (1997) Metabolism of hydrogen peroxide by the scavenging system in *Chlamydomonas reinhardtii*. *Physiol. Plant.* **99**, 49–55.

Takeda, T., Yoshimura, K., Yoshii, M., Kanahoshi, H., Miyasaka, H. and Shigeoka, S. (2000) Molecular characterization and physiological role of ascorbate peroxidase from halotolerant *Chlamydomonas sp* W80 strain. *Arch. Biochem. Biophys.* **376**, 82–90.

Toepel, J., Albaum, S.P., Arvidsson, S., Goesmann, A., la Russa, M., Rogge, K. and Kruse, O. (2011) Construction and evaluation of a whole genome microarray of *Chlamydomonas reinhardtii*. *BMC Genomics*, **12**, 579.

Tolleter, D., Ghysels, B., Alric, J., Petroutsos, D., Tolstygina, I., Krawietz, D., Happe, T., Auroy, P., Adriano, J.M., Beyly, A., Cuine, S., Plet, J., Reiter, I.M., Genty, B., Cournac, L., Hippler, M. and Peltier, G. (2011) Control of hydrogen photoproduction by the proton gradient generated by cyclic electron flow in *Chlamydomonas reinhardtii*. *Plant Cell*, **23**, 2619–2630.

Torzillo, G., Scoma, A., Faraloni, C., Ena, A. and Johanningmeier, U. (2009) Increased hydrogen photoproduction by means of a sulfur-deprived *Chlamydomonas reinhardtii* D1 protein mutant. *Int. J. Hydrogen Energy*, **34**, 4529–4536.

Urzica, E.I., Adler, L.N., Page, M.D., Linster, C.L., Arbing, M.A., Casero, D., Pellegrini, M., Merchant, S.S. and Clarke, S.G. (2012) Impact of oxidative stress on ascorbate biosynthesis in *Chlamydomonas* via regulation of the VTC2 gene encoding a GDP-L-galactose phosphorylase. *J. Biol. Chem.* **287**, 14234–14245.

Wang, Y., Ghaffari, N., Johnson, C.D., Braga-Neto, U.M., Wang, H., Chen, R. and Zhou, H. (2011) Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics*, **12**(Suppl 10), S5.

Wang, Z., Xu, C. and Benning, C. (2012) TGD4 involved in endoplasmic reticulum-to-chloroplast lipid trafficking is a phosphatidic acid binding protein. *Plant J.* **70**, 614–623.

Winkler, M., Hemschemeier, A., Jacobs, J., Stripp, S. and Happe, T. (2010) Multiple ferredoxin isoforms in *Chlamydomonas reinhardtii* – their role under stress conditions and biotechnological implications. *Eur. J. Cell Biol.* **89**, 998–1004.

Xu, C., Moellering, E.R., Muthan, B., Fan, J. and Benning, C. (2010) Lipid transport mediated by *Arabidopsis* TGD proteins is unidirectional from the endoplasmic reticulum to the plastid. *Plant Cell Physiol.* **51**, 1019–1028.

Zhang, Y., Adams, I.P. and Ratledge, C. (2007) Malic enzyme: the controlling activity for lipid production? Overexpression of malic enzyme in *Mucor circinelloides* leads to a 2.5-fold increase in lipid accumulation. *Microbiology*, **153**, 2013–2025.

Supporting information

Additional Supporting information may be found in the online version of this article:

Figure S1 (a) Histograms to test the normal distribution of differential expressed genes in *Chlamydomonas reinhardtii* during hydrogen production. (b) RT-PCR data of differential expressed genes in *Chlamydomonas reinhardtii* during hydrogen production.

Figure S2 RT-PCR data of differential expressed genes in *C. reinhardtii* during hydrogen production. Plotted are the mean values for time point T 3 (96 h) against time point t0 (0 h). For details and primers see Nguyen *et al.*, (2011).

Table S1 Differential expressed gene models in *Chlamydomonas reinhardtii* during hydrogen production determined with microarray.

Table S2 Differential expressed gene models in *Chlamydomonas reinhardtii* during hydrogen production determined with RNA seq.

Table S3 Time resolved cluster analysis of RNA-seq and microarray based transcript expression level (log2 ratio) for *Chlamydomonas reinhardtii* during hydrogen production (cluster were generated with GENESIS software, Pearson un-centered distance).

Table S4 Differential expressed gene models in *Chlamydomonas reinhardtii* during hydrogen production determined with RNA seq and microarray.

