Transcriptome analysis of the sulfate deficiency response in the marine microalga *Emiliania huxleyi*

Michal Bochenek¹,², Graham J. Etherington³, Anna Koprivova¹, Sam T. Mugford¹, Thomas G. Bell²,⁴, Gill Malin², and Stanislav Kopriva¹*

¹John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, UK

²Laboratory for Global Marine and Atmospheric Chemistry, School of Environmental Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK

³The Sainsbury Laboratory, Norwich NR4 7UH, UK

⁴Present address: Department of Earth System Science, University of California, Irvine, Croul Hall, Irvine, CA, 92697-3100, USA.

**Corresponding author**

Stanislav Kopriva

John Innes Centre

Norwich, NR4 7UH

UK

Tel. +44-1603-450276

fax +44-1603-450014
Word count:

Summary 198

Introduction 473

Materials and Methods 1465

Results and Discussion 4848

Total (Introduction, Materials & Methods, Results and Discussion, Acknowledgements) 6923

7 Figures, Fig. 6 and Fig. 7 coloured

1 Table
SUMMARY

• The response to sulfate deficiency of plants and freshwater green algae has been extensively analysed by system biology approaches. In contrast, seawater sulfate concentration is high and very little is known about sulfur metabolism of marine organisms.

• We used a combination of metabolite analysis and transcriptomics to analyse the response of the marine microalga *Emiliania huxleyi* as it acclimated to sulfate limitation.

• Lowering sulfate availability in artificial seawater from 25 mM to 5 mM resulted in significant reduction in growth and intracellular concentrations of dimethylsuloniopropionate and glutathione. Sulfate limited *E. huxleyi* cells showed increased sulfate uptake but sulfate reduction to sulfite did not seem to be regulated.

• Sulfate limitation in *E. huxleyi* affected expression of 1,718 genes. The vast majority of these genes were up-regulated, including genes involved in carbohydrate and lipid metabolism and genes involved in the general stress response.

• The acclimation response of *E. huxleyi* to sulfate deficiency shows some similarities to the well described responses of *Arabidopsis* and *Chlamydomonas* but also has many unique features. This dataset shows that even though *E. huxleyi* is adapted to
constitutively high sulfate concentration it retains the ability to re-program its gene expression in response to reduced sulfate availability.

**Keywords**

Sulfate deficiency, sulfur metabolism, transcriptomics. RNA sequencing, *Emiliania huxleyi*, DMSP, dimethylsulfoniopropionate, glutathione
Sulfur is essential for growth of all living organisms. Photosynthetic organisms cover their sulfur requirements by taking up and assimilating inorganic sulfate (reviewed in Takahashi et al., 2011a). In terrestrial and freshwater ecosystems, sulfate is often in low concentration and limiting growth, therefore, specific responses to sulfate deficiency have evolved in various taxa. Plants and freshwater algae respond to sulfate deficiency primarily by increasing sulfate uptake and assimilation capacity (Yildiz et al., 1994; Clarkson et al., 1999; Hoefgen & Nikiforova, 2008), while, e.g., the green alga *Chlamydomonas reinhardtii* is also capable of inducing arylsulfatases needed for utilisation of alternative sulfur sources (de Hostos et al., 1988). Sustained sulfate deficiency leads to decrease in photosynthetic activity and to reprogramming the plant metabolism to ensure sufficient resources are available for seed production (Hoefgen & Nikiforova, 2008).

In contrast, sulfate is plentiful for the diverse sulfate reducing organisms in the sea, since its concentrations in seawater reaches 25-28 mM. The high sulfate concentration in marine environment seems to be an evolutionary driver in expansion of modern phytoplankton groups (Ratti et al., 2011). In accordance with the high availability of sulfate in seawater, many phytoplankton species synthesise and accumulate large amounts of sulfur-containing metabolite, dimethylsulfoniopropionate (DMSP). This compound has been proposed many roles in phytoplankton, from a simple osmolyte or a sink for excess electrons to a signal
molecule for biotic interactions in marine environment (Steffels, 2000; Steinke et al., 2006).

It is also a precursor of dimethylsulfide (DMS), an atmospheric gas with a great impact on geochemical sulfur cycle and possibly climate (reviewed in Giordano et al., 2005; Quinn & Bates, 2011).

Because of the contrasting sulfate concentration in marine environment (25-28 mM) and soil/freshwater (10-50 µM), sulfate has never been considered to limit productivity in the oceans. We, however, hypothesised that phytoplankton is well adapted to the high sulfate availability so that reduction in its availability would affect growth. The intriguing question arising from this hypothesis is, whether marine phytoplankton retained the ability to respond to sulfate limitation, which is so prominent in plants and freshwater algae. The recent progress in genomics of marine phytoplankton (Tirichine & Bowler, 2011) paved a way for addressing this question on an overall transcriptome level. Among the phytoplankton species available for such analysis, the coccolithophore *Emiliania huxleyi* seems to be the most suitable model because of the large intracellular contents of DMSP, which accumulates to 50–242 mM in different strains of *E. huxleyi* (Steinke et al., 1998).

Here we show that lowering sulfate availability to concentrations below those that *E. huxleyi* encounters in the natural environment indeed reduces growth and DMSP concentration.

Transcriptomics analysis on *E. huxleyi* as it adjusted to sulfate deficiency revealed that some
acclimation responses to this stress condition are conserved among *E. huxleyi*, *Chlamydomonas*, and *Arabidopsis*, but many are unique to this marine microalga.

MATERIALS AND METHODS

Algal material and growth conditions

*Emiliania huxleyi* CCMP1516 was obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (NCMA, Bigelow, Maine, USA). The alga was grown in axenic batch culture in 250 mL conical flasks with 150 mL artificial seawater medium (ESAW) (Berges *et al.*, 2001) in a growth chamber (MLR 351; Sanyo; Loughborough, UK) at 15°C under a light:dark cycle of 14:10 hours and an irradiance of 180 µE m⁻² s⁻¹. Culture flasks were gently stirred by hand on a daily basis. Based on frequent microscopy observation, this strain does not appear to produce coccoliths under these culture conditions. ESAW medium contains 25 mM Na₂SO₄ and 363 mM NaCl and for sulfate deficiency studies, the sulfate was replaced by NaCl to keep the ionic strength constant (i.e., each mmol Na₂SO₄ was substituted by 3 mmol NaCl). Three independent cultures per treatment were inoculated with the same volume of a control culture (50 µL, ca. 50,000 cells), which was grown in ESAW medium to the mid-logarithmic phase of growth. Samples for analysis were always taken about 2 hours into the light phase. For the transcriptome analysis the
samples were taken six days after inoculation, in the middle of the exponential phase, to ensure the cells were still in the process of acclimation.

**Determination of E. huxleyi growth**

Cell density (cells mL$^{-1}$) and cell volume ($\mu$m$^3$ mL$^{-1}$) were determined using a Coulter multisizer (Beckman Multisizer 3, High Wycombe, UK) with a 100 $\mu$m aperture tube. Measurements were done with 100 $\mu$L culture samples diluted to 10 mL with 0.2 $\mu$m filtered seawater. The growth rates were calculated as $\frac{\ln(N_2) - \ln(N_1)}{(d_2-d_1)}$, where $N_i$ is the number of cells at day $d_i$. The ratio of variable to maximum chlorophyll fluorescence (Fv/Fm) which gives an estimate of PS II efficiency of was measured with a Walz Phyto-Pam phytoplankton analyser (Heinz Walz GmbH, Effeltrich, Germany) after 30 min dark adaptation.

**Determination of DMSP**

*E. huxleyi* particulate DMSP (DMSPp) concentration in was determined using headspace gas chromatography (Steinke *et al.*, 2000). Two to three mL of culture was gently filtered through 25 mm Whatman GF/F filters (nominal pore size 0.7 $\mu$m) using a hand-operated vacuum pump. The filters were placed in 4 mL vials containing 3 mL of 0.5 M NaOH and the vials were immediately sealed gas-tight with a screw thread cap and Teflon coated...
To ensure complete cold alkali hydrolysis of DMSP to DMS vials were incubated for at least 24 h in darkness at room temperature. They were equilibrated for 1 h at 30°C before analysis using a gas chromatograph with a flame photometric detector (Shimadzu 2010, Milton Keynes, UK) and a 30 m x 0.53 mm CP SIL 5CB column (Varian, Wokingham, UK). For analysis, 50 µL of headspace gas was withdrawn using an autosampler (MPS 2, Gerstel, Mülheim, Germany) with a 100 µL gas-tight syringe and injected into the GC (Steinke et al., 2000). DMSP concentration was determined using a calibration curve for known quantities of DMSP (linear between 0.1 – 25 µM) treated by the same alkali hydrolysis procedure.

Adenosine 5’-phosphosulfate reductase activity

Adenosine 5’-phosphosulfate reductase (APR) activity was determined as the production of $[^{35}\text{S}]$sulfite, assayed as acid volatile radioactivity formed in the presence of $[^{35}\text{S}]$APS and dithioerythritol (Koprivova et al., 2008). Ten mL culture aliquots were centrifuged (10 min, 10,000 g), supernatants removed and pellets re-suspended in 1.5 mL of culture medium and re-centrifuged (5 min, 10,000 g). The cells were disrupted by sonication on ice in 500 µL extraction buffer (50 mM Na/KPO$_4$ pH 8; 30 mM Na$_2$SO$_3$; 0.5 mM AMP, 10 mM dithioerythritol). The extracts were centrifuged (30 sec, 1,000 g) to remove cell debris and 20 µL was used for APR measurement. Protein concentrations were determined with a
protein assay kit (Bio-Rad, Hemel Hempstead, UK), using bovine serum albumin as the standard.

HPLC analysis of low molecular weight thiols

Thiols were extracted from cells filtered from 15-25 mL culture aliquots using hot methanesulphonic acid (Dupont et al., 2004). Total cysteine and glutathione were analysed following the method of Koprivova et al., (2008).

Sulfate uptake

To measure the sulfate uptake, *E. huxleyi* cultures were grown in 500 mL conical flasks with 250 mL ESAW medium containing 25 mM (control) or 5 mM sulfate. Fifty mL of control or 100 mL of sulfate deficient cultures were filtered onto 47-mm diameter 1.2 µm filters (Millipore; Watford, UK), and washed with 200 mL S-free medium to remove sulfate. The cells were re-suspended in 50 mL tubes with 10 mL ESAW medium containing 25 mM or 5 mM sulfate. The cell density and volume was determined for each tube. $[^{35}\text{S}]$sulfate was added to a specific activity of 192 kBq mL$^{-1}$ and the cells were incubated for 60 min in the light. The cells were collected by filtration, washed twice with 100 mL S-free medium and placed into 20 mL scintillation vials. To dissolve the filters and disrupt the cells, 5 mL of tissue solubiliser (Solene®-350, PerkinElmer, Cambridge, UK) was added and the vials
were kept overnight at room temperature. The next day, 10 mL of scintillation cocktail Optisafe 3 (Perkin Elmer) was added and $[^{35}\text{S}]$ radioactivity was determined by scintillation counting (Wallac 1409, Perkin Elmer).

RNA Isolation and Expression Analysis

Total RNA was isolated by standard phenol/chlorophorm extraction and LiCl precipitation. For quantitative RT-PCR (qPCR) first-strand cDNA was synthesized from 1 µg of total RNA using a QuantiTect Reverse Transcription Kit (Qiagen, Crawley, UK), which includes a DNAse step. The qPCR reactions were performed in duplicate for 3 independent samples using gene-specific primers (Table S7) as described in Lee et al., (2011). All quantifications were normalized to the ACTIN gene (gene ID 74049).

For sequencing, total RNA from three independent sulfate limited and control cultures was repurified using an RNeasy Plant isolation kit (Qiagen) with on column DNAse removal of contaminating DNA. To prepare Illumina RNAseq libraries poly-A RNA was isolated from 5 µg of total RNA and the mRNA was fragmented to an average size of 100 bp using the manufacturer’s instructions (Illumina mRNAseq Kit, Illumina, Cambridge, UK). First strand cDNA synthesis used Superscript III reverse transcriptase (Invitrogen, Paisley, UK) and 3 µg random hexamer primers (Illumina). Second strand cDNA and RNAseq libraries were prepared according to the manufacturer’s instructions (Illumina). Following a PCR
amplification and size selection, the mRNA libraries were sequenced using the Illumina HiSeq 2000 platform to generate paired end 50 bp reads. All sequenced reads were quality controlled (QC) by removing reads containing ‘N’ s and those with read-lengths shorter or longer than 50 nucleotides. The QC reads for each biological replicate were separately aligned to the *E. huxleyi* reference transcriptome at JGI (Emihu1_best_transcripts.fasta from http://genome.jgi-psf.org/Emihu1/Emihu1.download.html) using TopHat (Trapnell et al., 2009, Trapnell et al., 2012). Further analyses were performed with the Cufflinks tools suite (Trapnell et al., 2010). Differential expression between the control and sulfate deficient cultures was calculated by Cuffdiff using the FPKM (Fragments Per Kilobase of exon per Million fragments mapped) normalisation, false discovery rate of 5%, and Cuffdiff-min-alignment-count parameter of 622. The sequenced reads were then aligned to the *E. huxleyi* reference genome (Emihu1_scaffolds.fasta) using TopHat to examine the number of reads that mapped to the genome but not the transcriptome. Differential expression was then calculated using the Tophat-Cufflinks-Cuffcompare-Cuffdiff pipeline. To compare our transcripts to the predicted *E. huxleyi* transcripts, the Cuffdiff transcript expression file (providing the assembled transcripts) was compared to that of the Emihu_1_best_genes.gff annotation file from JGI. Reads that did not map to the *E. huxleyi* reference genome were de novo assembled using OasesOptimizer. The resulting transcript assembly was used as a reference and the unmapped
reads were analysed for differential expression. To obtain insight into the identity of the novel transcripts, they were subjected to BLAST analysis (for details of the procedures see Supplementary methods).

**Functional annotation**

Superfamily information for the *E. huxleyi* transcripts was obtained from Superfamily database 1.73 (http://supfam.cs.bris.ac.uk/SUPERFAMILY/cgi-bin/gen_list.cgi?genome=ex). To identify biological processes affected by sulfate deficiency all transcript data were ranked according to expression ratio and analysed by iterative group analysis (Breitling *et al.*, 2004). Functional categories of the KEGG (Ogata *et al.*, 1999) and KOG (Tatusov *et al.*, 2003) database were downloaded from http://genome.jgi-psf.org/Emihu1/Emihu1.download.ftp.html. Genes encoding sulfate transporters and components of sulfate assimilation were identified using BLAST at the JGI site, and the identities of genes discussed were confirmed by BLAST at the NCBI site (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

For comparison with other organisms, transcriptomics data from *A. thaliana* grown for 8 days on normal sulfate supply and transferred on S limited conditions for 6 days (Nikiforova *et al.*, 2003) and *Chlamydomonas* exposed to S starvation for 6 h (González-Ballester *et al.*, 2010) were used. In these studies *A. thaliana* and especially
Chlamydomonas would still have been acclimating to S limitation, thereby representing a good comparison with our data for *E. huxleyi*.

RESULTS AND DISCUSSION

*E. huxleyi* growth is limited by low sulfate concentration

Growth of plants and green algae can be limited by sulfur limitation, therefore, we tested whether sulfate availability alters *E. huxleyi* growth. Batch cultures of *E. huxleyi* were grown at sulfate concentrations of 25 mM (control), 10 mM, 5 mM and 1 mM, and cell density and volumes were monitored (Figure 1a and b). Lowering the sulfate concentration to 10 mM had no effect on growth rate. Interestingly, the cell volume reduced by ca. 60% over the time course of the experiment with 10 and 25 mM sulfate; similar decreases in cell volume have been shown for other strains of *E. huxleyi* (van Rijssel & Gieskes, 2002).

Further reduction to 5 mM sulfate reduced the growth rate in the exponential phase by approximately 50% from 0.62 in the controls to 0.30 d\(^{-1}\) and the cells were also about two-fold larger than those grown at higher sulfate concentrations (Figure 1a and b). This sulfate concentration is 3 orders of magnitude higher than the concentration that limits the growth of plants; indeed, plants grow normally even at 5 µM sulfate (Hawkesford and De Kok, 2006), while freshwater algae adapt to environments with sulfate concentrations in the 0.01-1 mM range (Giordano *et al.*, 2005). At 1 mM sulfate growth was very slow (0.05 d\(^{-1}\);
Figure 1a) and cell volume was ca. six-fold larger volume than control cells by day 13 (Figure 1b). The efficiency of PS II was not substantially affected by low sulfate conditions apart from an initial drop in Fv\Fm in the 1 mM cultures that within two days recovered to control levels (Figure S1).

These results are consistent with a recent report by Ratti et al., (2011) showing significantly slower growth rates for *E. huxleyi* (strain PML92/11) at 5 mM and 1 mM sulfate compared to 10 mM and higher sulfate concentrations. The same was true for the dinoflagellate *Protoceratium reticulatum* (PRA0206), but not for the green alga *Tetraselmis suecica* (PCC 305) or the marine cyanobacterium *Synechococcus* sp. (UTEX LB2380) (Ratti et al., 2011).

Thus it seems that increase in sulfate concentration in sea water was one of the major evolutionary drivers for the success of chlorophyll a+c phytoplankton, including the coccolithophores (Ratti et al., 2011). It is interesting to note that the *E. huxleyi* cultures grew normally at 10 mM sulfate, as this was the concentration in water in which these algae evolved (Ratti et al., 2011). Also the growth of the halophilic green alga *Dunaliella salina* was limited by sulfate concentrations lower than 0.1 mM (Giordano et al., 2000). The cessation of cell division and increase in cell size observed in *E. huxleyi* at 1 mM sulfate was similar to the effects of sulfur deprivation on *Chlamydomonas* (Melis et al., 2000; Zhang et al., 2002), but may also be a response to uncoupling between growth rate and division rate.
Decreased sulfate availability also affected intracellular DMSP concentration (Figure 1c). Under control conditions intracellular DMSP remained at a stable and high level (ca. 260 mM) throughout the experiment, whereas cells growing at lower sulfate concentrations showed a notable reduction in DMSP concentration. DMSP decreased until day 8, and then remained relatively stable at ca. 75%, 40%, and 20% of control levels for cultures containing 10 mM, 5 mM and 1 mM sulfate, respectively. Interestingly, at 10 mM sulfate the DMSP concentration was altered but growth was unaffected, whereas cultivation at 5 mM and 1 mM sulfate decreased both growth and DMSP accumulation. Surprisingly, Ratti et al., (2011) did not observe a significant change in DMSP content in *E. huxleyi* strain PML92/11 between sulfate concentrations of 5 mM and 20 mM. However, the maximal intracellular DMSP concentration in the strain PML92/11 reached only ca. 100 mM (Ratti et al., 2011), which is similar to the concentration found in *E. huxleyi* CCMP 1516 grown at 1 mM sulfate (Fig. 1). Given that a five-fold reduction in sulfate concentration in the medium resulted in reduced growth rate and DMSP accumulation, 5 mM sulfate was chosen as the sulfate deficiency treatment for all subsequent experiments. It has to be noted, however, that our experiments used sulfate deficiency as a tool rather than to mimic environmental conditions.

Next we tested whether the observed changes in growth were reversible and thereby connected to sulfate availability. *E. huxleyi* was grown in 25 mM and 5 mM sulfate media
for 8 days, i.e. late exponential phase. Half of the sulfate deficient cultures were supplemented with sulfate to restore the concentration to the control 25 mM level. The sulfate addition rapidly increased the specific growth rate from 0.15 to 0.32 d\(^{-1}\) and after 3 days the cell numbers in the supplemented cultures were only 10% lower than in the controls and double that of the sulfate deficient cultures (Figure 2a). This increase in cell numbers was accompanied by a decrease in cell volume which reached control levels after 48 h (Figure 2b). Sulfate restoration also enhanced DMSP accumulation and after 48 h DMSP levels were indistinguishable from those of the control cells (Figure 2c). One of the markers for sulfate deficiency in plants is the induction of APR, therefore, we determined the activity of this enzyme in the three treatments. Unlike intracellular DMSP concentration, APR activity was not significantly different in the different cultures throughout the experiment. The activity decreased in all the cultures at 24 h and thereafter, but this was probably connected with the end of the exponential phase of growth (Figure 2d).

Regulation of *E. huxleyi* sulfur metabolism by sulfate deficiency

Having determined that a decrease in sulfate concentration from 25 mM to 5 mM limits growth and DMSP accumulation in *E. huxleyi*, we examined the acclimation response of the algae to these conditions. Cellular cysteine content decreased by ca. 60% during culture growth, but no difference was detected between sulfate deficient and control cultures.
(Figure 3a). However, as with DMSP, glutathione (GSH) concentration was about two-fold lower in sulfate deficient cells than in the controls after 6 and 10 days of cultivation (Figure 3b). A reduction in GSH concentration is a typical consequence of sulfate deficiency in plants (Hirai et al., 2003, Nikiforova et al., 2003). In plants and Chlamydomonas, another typical response to sulfate starvation is an increase in sulfate uptake capacity (Pootakham et al., 2010). We tested E. huxleyi cultures grown at 5 mM and 25 mM sulfate for $^{35}$S sulfate uptake at both concentrations and found that the 5 mM grown cells showed an approximately three-fold higher uptake of sulfate at both concentrations compared to control cells grown at 25 mM sulfate (Figure 4). Thus, in common with plants, E. huxleyi reacts to sulfate deficiency by increasing sulfate uptake.

General transcriptome analysis of E. huxleyi response to sulfate deficiency

To enable comparison of the acclimation responses to sulfate deficiency of E. huxleyi with other organisms, we analysed the transcriptomes of sulfate deficient and control E. huxleyi cells. Total RNA was isolated from three independent sulfate deficient and control cultures. At the time of sampling the average growth rates were 0.67 day$^{-1}$ and 0.46 day$^{-1}$ in the control and sulfate deficient cultures, respectively. The control cultures achieved 5 generations from the beginning of the treatment, whereas the sulfate deficient cultures 4 generations, reaching average intracellular DMSP concentrations of 226 and 103 mM,
respectively. The RNA was subjected to RNA sequencing using the Illumina platform at The GenePool Edinburgh. The resulting 50 bp reads for each biological replicate were aligned separately to the *E. huxleyi* CCMP1516 reference transcriptome (Emihu1_best_transcripts.fasta) based on the Joint Genome Institute (JGI). The *E. huxleyi* genome assembly predicted 39,125 genes, compared with predicted 10,402 gene models in *Phaeodactylum tricornutum* (Bowler et al., 2008) and 11,390 models in *Thalassiosira pseudonana* (Armbrust et al., 2004). The much larger number of gene models is due to the presence of diploid alleles for many genes.

From a total of 58,871,530 reads (after quality control), 37,008,141 mapped to the reference transcriptome, leaving 21,863,389 (37.1%) reads unmapped (Figure 5, Table 1). From the 39,125 predicted *E. huxleyi* transcripts in the JGI reference transcriptome, 16,729 transcripts were identified using the Cufflinks tool suite. The expression levels were normalised using the FPKM as a measure of expression strength. Differential expression was then calculated for each locus shared between the control (ehux25) and sulfate deficient (ehux5) samples using a false discovery rate of 5%. Amongst the shared Cufflinks transcripts 278 were found to be differentially expressed (Table S1) using the Cuffdiff min-alignment-count parameter as a threshold (see Supplementary methods). Of these transcripts 224 were up-regulated and 54 were down-regulated in the sulfate deficient cultures. In addition, 29 *E.
huxleyi transcripts were only expressed in the ehux25 dataset and 1029 E. huxleyi
transcripts were expressed solely in ehux5 (Table S2).

Given the large number of reads unmapped to the transcriptome, we mapped all reads to the
E. huxleyi reference genome (Emihu1_scaffolds.fasta) to identify previously undiscovered
transcripts. Using this approach, we found that 48,367,452 reads mapped to the reference
genome, leaving 10,504,078 (17.78%) of the reads unmapped (Table 1). Through this
mapping we identified 20,416 transcribed genomic loci (Figure 5) and compared them to
the 39,125 predicted transcripts. The total transcript space for our models was 42,331,395
nucleotides with an average transcript-length of 2,073 bp (compared to 67,258,384 bp in
Emihu1_best_transcripts, with an average transcript-length of 1,719 bp). Of our assembled
transcripts, 15,680,463 nucleotides had no overlap with the E. huxleyi predicted transcripts
(i.e. they were ‘missing’ from the E. huxleyi annotated gene models). This assembly
resulted in 11,576 novel transcripts (not overlapping any of the predicted E. huxleyi
transcripts) and 8,840 transcripts that overlapped with the annotated transcripts. The amount
of overlap between our newly assembled transcripts and the predicted gene models at the
JGI site was 46.15%. Qualitatively, this is consistent with analysis of general gene
expression in P. tricornutum by EST sequencing which identified an additional 1,968
transcripts not predicted as gene models (Maheswari et al., 2010), and the 3,470
unpredicted transcripts in T. pseudonana tiling array data (Mock et al., 2008). The large
number of unannotated transcripts is however surprising and may reflect the difficulties of
the gene prediction software regarding the high GC content of *E. huxleyi* genome. Thus, the
*E. huxleyi* genome and transcriptome are much larger than those of the two sequenced
diatom species. The calculation of differential expression of these transcripts determined by
Cuffdiff resulted in identification of 325 loci, 254 up-regulated and 71 down-regulated (Table S3). Among the up-regulated and down-regulated transcripts that mapped to genome
scaffolds 121 and 31 transcripts, respectively, did not overlap with any gene models (Table
S3).

The 10,504,078 reads that would not map to the reference genome (3,467,678 paired-end
reads and 3,568,722 single reads) were assembled de novo. Using the Cufflinks tool suite,
7,712 expressed transcripts were identified. Differential expression was then calculated for
each locus shared between ehux25 and ehux5. From this we found 335 transcripts that were
significantly differentially expressed: 298 up-regulated and 37 down-regulated in ehux5
(Table S4). The identity of these new transcripts was investigated by BLAST. Some of
these unmapped transcripts correspond to chloroplast and mitochondrial transcripts that are
not included in the scaffolds, however, the majority of the most highly up-regulated genes
had no homology with any known sequences. Altogether, the analysis identified 1,718
differentially regulated transcripts (325 from genome alignment, 1058 present in one
condition only, 335 novel transcripts), 1,253 of them up-regulated.
The expression pattern obtained by RNA sequencing was verified by quantitative RT-PCR (qPCR) for seven differentially regulated genes with gene IDs: 450514, 452597, 315901, 229382, 456731 (up-regulated) and 454260 and 432295 (down-regulated). For all these genes the difference in transcript abundance according to RNA sequencing data agreed well with the qPCR results (Figure S2).

Functional categories of genes affected by sulfate limitation in *E. huxleyi*

To identify functional categories, the transcripts were annotated using the Superfamily database 1.73 and this enabled description of 14,249 genes for which domain assignment was available. To obtain information on biological processes affected by sulfate deficiency the genes were further characterised by KEGG and KOG annotation (Ogata *et al.*, 1999; Tatusov *et al.*, 2003). Figure 6 shows the distribution of KOG functional categories amongst the transcripts up-regulated by sulfate deficiency. The most prominent functional groups were “Signal transduction mechanisms” and “Post-translational modification, protein turnover, chaperones”.

To obtain a deeper insight into the biological processes affected by sulfate deficiency iterative Gene Analysis (iGA) was employed. Using the KEGG pathway classification 2,723 genes were annotated and assigned to 90 metabolic pathways. The iGA identified 23
up-regulated and 30 down-regulated KEGG pathways in the transcriptome of sulfate
deficient cells (Table S5). Amongst the up-regulated pathways 4 can be assigned to
carbohydrate metabolism (ascorbate and aldarate, butanoate, and pyruvate metabolism, and
citrate cycle) and 5 pathways to lipid metabolism (fatty acid, bile acid, sphingoglycolipid,
prostaglandin and leukotriene, and glycerolipid metabolism). By contrast, the down-
regulated gene set had no single prominent group but is rather a representative cross section
of various KEGG classes. Among the KEGG pathways affected by sulfate deficiency,
several linked to sulfur metabolism were found. Glutathione metabolism was found in the
up-regulated pathways, whereas methionine, thiamine and serine metabolic pathways were
down-regulated. Interestingly, degradation pathways for a range of xenobiotics were down-
regulated under sulfate deficiency, likely reflecting the reduced availability of GSH.
The iGA analysis was also performed using KOG annotation of 9,725 transcripts clustered
into 2,254 KOG functional groups. Table S5 shows the 40 and 23 KOG groups enriched
amongst the up- and down-regulated transcripts, respectively. The KOG groups uniquely
detected in the up-regulated list included 10 clusters assigned to the “Cytoskeleton” class, 5
clusters to “Inorganic ion transport and metabolism” and single clusters assigned to “Energy
production and conversion”, “Nuclear structure” and “Function unknown”. In contrast, the
groups appeared more evenly distributed throughout the down-regulated list. Importantly,
sulfate deficiency resulted in an enrichment of the “Sulfate/bicarbonate/oxalate exchanger
SAT-1 and related transporters (SLC26 family)” class, driven mainly by strong induction of two sulfate transporter transcripts (441761, 453061), which agrees well with the increased sulfate uptake capacity of sulfate deficient *E. huxleyi* cells (Figure 4).

**Transcriptional regulation of enzymes of sulfur metabolism**

Given the known effects of sulfate deficiency on sulfate uptake and assimilation in various organisms, we compared the expression of genes involved in these processes in sulfate deficient and control *E. huxleyi* cells. It was shown previously that the sulfate assimilation pathway in *E. huxleyi* is organised in a similar way to plants (Kopriva *et al.*, 2009) and the corresponding genes have been identified (Table S6).

**Sulfate transport.** Our BLAST analysis identified 16 putative sulfate transporters (STR) in the haploid *E. huxleyi* genome (Table S6, Figure S3). The transporters fall broadly into 4 groups, the plant/fungi/animal SCL26 type, the SAC1/SLT Na\(^+\)/SO\(_4^{2-}\) co-transporter family, the SLC13 family, and several other transporters characterised by the sulfate transporter and anti-sigma antagonist (STAS) domain (Takahashi *et al.*, 2011b). In accordance with the increased sulfate uptake (Figure 4), transcripts of 3 putative sulfate transporters *STR1* (protein ID 363809), *STR2* (441761), and *STR3* (453061) increased in sulfate-limited cells 5-, 10-, and 15- fold, respectively (Figure 7, Table S6). These genes encode transporters of the SLC26 family and contain the characteristic STAS domain (Takahashi *et al.*, 2011b).
and therefore probably represent real sulfate transporters. In addition, one gene (*STR13*; 230466) from the group of STAS containing genes and an SLC13 family gene (*STR16*; 443760) have been up-regulated. This agrees well with the induction of sulfate transporters in *Chlamydomonas* (González-Ballester et al., 2010) and *Arabidopsis* (Maruyama-Nakashita et al., 2006) and is shown for comparison in Figure 7. However, it is highly probable that not all the 16 *STR* genes encode genuine sulfate transporters. The SLC13 transporters often function as Na⁺ di-or tricarboxylate carriers and may not participate in sulfate uptake in *E. huxleyi*. The genes of the SAC1/SLT family are only distantly related to their *Chlamydomonas* counterparts and their role in sulfate transport has yet to be established, particularly as they were not regulated by sulfate deficiency in *E. huxleyi*, whereas the *Chlamydomonas SLT1* and *SLT2* are up-regulated by sulfate starvation (Pootakham et al., 2010).

**Sulfate reduction.** A common response to sulfate deficiency is the increase in sulfate reducing capacity by up-regulation of APR in plants (Nikiforova et al., 2003), or ATP sulfurylase and sulfite reductase in *Chlamydomonas* (Ravina et al., 2002). Surprisingly, however, in *E. huxleyi* neither APR activity nor transcript levels for *APR, ATPS*, and *SiR* were affected by sulfate deficiency (Figure 7). Interestingly, while APR activity in *E. huxleyi* was not regulated by changes in sulfate levels, it was about ten-fold higher than APR activity typically measured in *Arabidopsis* or *Chlamydomonas* (Vauclare et al., 2002;
Ravina et al., 2002). APR activity and the general capacity to reduce sulfate might thus be high enough in *E. huxleyi*, such that further increases would be meaningless.

**Cysteine synthesis.** The transcript levels of genes involved in cysteine synthesis, *OASTL4* (452198), *OASTL6* (445218), *OASTL7* (430252), *OASTL8* (440100), and *OASTL10* (442172), isoforms of OAS thiolyase and *SAT3* (248485) and *SAT4* (234967) of serine acetyltransferase, increased under sulfate deficiency. Thus, in all three organisms, *E. huxleyi*, *A. thaliana*, and *C. reinhardtii*, at least one isoform of serine acetyltransferase (SAT) and OAS thiolyase were induced by sulfate deficiency (Fig. 7). In plants SAT has an important role in controlling the sulfate assimilation pathway: its overexpression increases the content of sulfur-containing metabolites and strong inhibition leads to growth reduction (Blaszczyk et al., 1999; Haas et al., 2008). The increase in SAT expression in *E. huxleyi* may facilitate cysteine synthesis when the substrate concentration is strongly diminished.

The regulation, or the lack thereof, of *STR1*, *STR2*, *STR3*, *ATPS1*, *APR*, *SAT3*, and *OASTL6* was confirmed by qPCR (Figure S2, Table S6).

**Glutathione metabolism.** GSH metabolism was found among the KEGG pathways up-regulated by sulfur limitation (Table S5). This was mainly because of a strong up-regulation of four genes encoding GSH-transferases (233986, 224152, 349113, 442908) and a GSH peroxidase (433534), which are connected with oxidative stress rather than GSH metabolism and reflect the general up-regulation of stress-related genes by sulfur deficiency.
However, two isoforms of GSH synthetase (51736, 121060) were also strongly, ca. six-fold, up-regulated as well as a minor isoform of γ-glutamylcysteine synthetase (113513) the first enzyme in GSH synthesis. This contrasts with no transcriptional regulation of GSH synthesis in *Arabidopsis* and *Chlamydomonas*. However, in *Arabidopsis* the γ-glutamylcysteine synthetase, which has much higher control over GSH synthesis, is regulated post-translationally by redox state (Hicks *et al.*, 2007). The genes for this enzyme in *E. huxleyi* have a different evolutionary origin, being more similar to animal genes than plant ones, and therefore the pathway may be regulated differently.

**Methionine metabolism.** In *E. huxleyi*, methionine is not only an essential amino acid for protein synthesis but also a precursor for DMSP synthesis. Interestingly, Met metabolism was among the KEGG pathways down-regulated by sulfur limitation. Indeed, the genes for two components of S-adenosylmethionine (SAM) cycle, SAM synthase and S-adenosylhomocysteine hydrolase were significantly down-regulated, by 15% and 40%, respectively (Table S6). On the other hand, homocysteine S-methyltransferase and cobalamin-independent methionine synthase, catalysing the last step of Met synthesis, were induced by sulfur deficiency. This response is more similar to *Chlamydomonas* than to *Arabidopsis*, since in the green alga the genes of SAM cycle were down-regulated while they were up-regulated by S deficiency in *Arabidopsis* (Nikiforova *et al.*, 2003, González-Ballester *et al.*, 2010). On the other hand, the Met metabolism is coordinately down-
regulated in *Chlamydomonas* (González-Ballester *et al.*, 2010), whereas some genes for Met synthesis were up-regulated in *E. huxleyi*. The genes of the alternative pathway of Met recycling, the Yang cycle, have been found in *E. huxleyi*, but expressed to very low levels, particularly compared to the SAM cycle (Table S6), so this pathway may not play a very important role. Interestingly, all five genes involved in the SAM cycle belong among the 60 most highly expressed genes in control cultures of *E. huxleyi*, as judged from the FPKM values, pointing to a very high importance of this pathway for the alga, connected most probably with high need for methylation, including DMSP synthesis.

**Sulfolipids.** In sulfate deficient *Arabidopsis* sulfolipid content is strongly reduced and the genes for sulfolipid synthesis are down-regulated (Nikiforova *et al.*, 2003). In contrast the genes are up-regulated in *C. reinhardtii* (González-Ballester *et al.*, 2010), reflecting the much larger sulfolipid pool in this alga, where sulfolipids are actively degraded as a source of sulfur for protein synthesis (Sugimoto *et al.*, 2007, 2010). In *E. huxleyi* sulfolipid synthesis genes were not differentially regulated, suggesting that sulfolipid turnover is not affected as in *Chlamydomonas*. This might reflect the importance of sulfolipids for marine organisms adapted to low phosphate availability (Van Mooy *et al.*, 2006) on one hand, and the presence of a large sulfur pool in DMSP suitable for sulfur recycling during sulfate limitation, on the other hand.
Arylsulfatases. *Chlamydomonas* responds to sulfate deficiency by induction of extracellular sulfatases that allow utilisation of organic sulfates (de Hostos *et al.*, 1988). These enzymes are not present in higher plants, but in *E. huxleyi* transcripts for arylsulfatases (95583, 107777, 433677) were found only in transcriptome of the sulfate deficient cells (Table S2) suggesting a similar mechanism for sulfur scavenging.

Signalling. While the response of gene expression and metabolite accumulation to sulfate deficiency has been well described, much less is known about the molecular mechanisms of sulfate-sensing and signalling. In *Arabidopsis*, the SLIM1 transcription factor is responsible for up-regulation of sulfate transporter genes (Maruyama-Nakashita *et al.*, 2006), whereas in *Chlamydomonas* the SAC1, Na\(^+\)/SO\(_4\)^{2-} transporters seem to be the sensor of sulfate status, and the SNRK2.1 and SNRK2.2 (SAC3) are essential for the transcriptional response (Davies *et al.*, 1999; González-Ballester *et al.*, 2008). In the *E. huxleyi* genome there are several genes belonging to the same family as SAC1. However, SAC1 itself is not regulated by sulfate starvation in *Chlamydomonas* (González-Ballester *et al.*, 2010), and the same is true for all the *E. huxleyi* genes of the SAC1/SLT group of transporters. Similarly, there are more than 90 genes with similarity to SAC3 in the *E. huxleyi* genome so it is impossible to assign a similar function to any of them. No protein homologous to SLIM1 is encoded in *E. huxleyi* genome.
General response to sulfate deficiency

The fundamental difference in the response to sulfate deficiency in *E. huxleyi* compared to *Arabidopsis* and *Chlamydomonas* is the ratio between up-regulated and down-regulated genes. The general response to prolonged sulfate deficiency in plants, equivalent to the late acclimation phase of the *E. huxleyi* cultures, is a slowing down of metabolism and shortening of the life cycle (Hoefgen & Nikiforova, 2008). Accordingly, in multiple microarray experiments significantly more transcripts were repressed by sulfate starvation rather than induced (Hirai *et al.*, 2003, Maruyama-Nakashita *et al.*, 2003; Nikiforova *et al.*, 2003). The same was true for *Chlamydomonas*, where greater than two-fold more transcripts were down-regulated by sulfate deficiency than up-regulated (González-Ballester *et al.*, 2010), and for *D. salina* where sulfate deficiency resulted in decreased Rubisco accumulation and PEP carboxylase and nitrate reductase activities (Giordano *et al.*, 2000).

In contrast, in *E. huxleyi* 1,029 transcripts were present only in sulfate deficient cells compared to 29 that were found only in the controls. Also among transcripts detected in both conditions, greater than 4-fold more genes were up-regulated than down-regulated. The up-regulated transcripts include many that reflect general stress including GSH transferases (see above), flavodoxin (ID 68288), protein disulphide isomerase (443239, 447219), or immunophillin (435425) (Table S1), all of which are also induced in sulfate deficient *Chlamydomonas* (González-Ballester *et al.*, 2010). Among other stress related
genes up-regulated in sulfur limited *E. huxleyi*, two encode GSH peroxidase (433534, 67177; Table S6). GSH peroxidase is particularly interesting, since in many organisms this enzyme contains selenocysteine (Forstrom *et al.*, 1978). *E. huxleyi* also synthesise selenoproteins, however, it is unique in primarily taking up selenite and not selenate (Araie & Shiraiwa, 2009). The two confirmed selenoproteins of *E. huxleyi* (443239, 417208) were also up-regulated by sulfur deficiency. Transcripts encoding several selenoproteins, and two Se binding proteins, accumulated in sulfur deprived *Chlamydomonas* cells while a gene encoding Se-binding protein was upregulated in *Arabidopsis* (Nikiforova *et al.*, 2003, González-Ballester *et al.*, 2010). This regulation in *Chlamydomonas* and *Arabidopsis* may be a response to increased uptake of selenate, caused by up-regulation of sulfate transporters, which are capable of uptake of selenate, but not selenite. Thus, the driver for up-regulation of the selenoproteins in *E. huxleyi* is more probably their function in stress response and not a sink for Se.

The decline in transcript levels for genes associated with photosynthetic electron transport, chlorophyll biosynthesis, and light harvesting observed in *Arabidopsis* and *Chlamydomonas* (Nikiforova *et al.*, 2003; González-Ballester *et al.*, 2010) were not observed in *E. huxleyi*. In accordance, our Fv/Fm data showing no effect of the limitation on Photosystem II quantum yield indicated that the substantial decrease in photosynthesis and chlorophyll synthesis observed in sulfate deficient plants and green algae (Wykoff *et al.*, 1998; Giordano *et al.*, 2009) was not associated with a decrease in the expression of genes involved in photosynthesis.
2000; Maruyama-Nakashita *et al.*, 2003; González-Ballester *et al.*, 2010) might possibly not be so dramatic in *E. huxleyi*. Photosynthesis has, however, a high demand for reduced sulfur to ensure synthesis of proteins and co-enzymes. It is possible that during the acclimation response to sulfate limitation *E. huxleyi* uses sulfur re-allocated from the large DMSP pool and so does not need to reduce their synthesis and limit photosynthesis.

The level of intracellular DMSP decreased concurrently with decreasing sulfate in the *E. huxleyi* cultures. Rather than a simple decrease in concentration due to reduced sulfate availability, this might be an active process to redirect sulfur from DMSP into other metabolic processes. This is corroborated by up-regulation of genes involved in the synthesis of alternative osmolytes proline and glycine betaine, pyrroline-5-carboxylate reductase (protein ID 349043) and betaine-aldehyde dehydrogenase (437142, 417844), respectively. Unfortunately, the genes involved in DMSP synthesis have not yet been unequivocally identified, so it is impossible to establish whether the decrease in DMSP is caused by the down-regulation of its synthesis. Lyon *et al.*, (2011) proposed 4 enzymes to catalyse DMSP synthesis based on their regulation by salinity in the diatom *Fragilariopsis cylindrus*. *E. huxleyi* homologues of 2 of these genes, aminotransferase (456646, 369841) and diaminopimelate decarboxylase (438904), were highly expressed as expected for a major pathway, but not significantly (q>0.05) regulated. The other 2 genes, S-

adenosylmethionine methyltransferase (464166, 254918) and NADPH reductase (100136,
106956, 120452) were either not expressed at all or expressed to a very low level and are thus very unlikely to participate in DMSP synthesis in this alga. However, in line with the sulfur redistribution hypothesis, (Stefels, 2000), a significant up-regulation of two genes encoding proteins with IDs 459683 and 470487 was observed. These genes, annotated as Class III acyl CoA transferases, are homologues of the bacterial $DddD$ (DMSP-dependent DMS production) genes involved in DMSP degradation to DMS and 3-hydroxypropionate (Todd *et al.*, 2007; 2010). DMSP synthesis is not only a large pool for sulfur, but also a significant sink for carbon. One of the proposed functions of DMSP is an overflow metabolite allowing safe dissipation of excess energy and reducing power (Stefels, 2000). In sulfate deficient *E. huxleyi* cells transcript levels for genes involved in the citric acid cycle, succinyl-CoA synthetase (417649), succinate dehydrogenase (432409), and citrate synthase (467883), and fatty acid biosynthesis, acetyl-CoA carboxylase (455280), beta-ketoacyl-ACP reductase (433820), or acyl-CoA dehydrogenase (437926) increased. It is thus possible to speculate that carbon that cannot be used for DMSP synthesis might be redirected into synthesis of fatty acids as was seen before for diatoms subjected to nitrogen deficiency (Hockin *et al.*, 2012).

Another process specific for *E. huxleyi*, which can be relevant for its response to sulfate starvation, is calcification. *E. huxleyi* can respond to the needs to dissipate excess energy, e.g. during high light intensities, by increasing the degree of calcification (Paasche, 2001;
Xu & Gao, 2012). Such similar physiological role of calcification and DMSP metabolism may explain the overall higher intracellular DMSP concentration and its notably greater decrease under low sulfate concentration in naked cells of *E. huxleyi* CCMP 1516 from this study compared to the calcifying strain PML92/11 from Ratti *et al.*, (2011). A series of experiments using calcifying strains subjected to various concentrations of sulfate, calcium and irradiance intensity would shed more light on the link between DMSP metabolism and calcification.

In conclusion, we have shown that, despite being adapted to high sulfate concentrations in sea water, the marine microalga *Emiliania huxleyi*, still retains the genetic program to respond to artificial sulfate deficiency. Whereas the up-regulation of sulfate uptake and cysteine synthesis in *E. huxleyi* is in common with plants and freshwater algae, the general response is significantly different. Instead of slowing down photosynthesis and primary metabolism *E. huxleyi* responds to sulfate deficiency by up-regulation of genes involved in carbohydrate and fatty acid synthesis and appears to redirect sulfur and carbon from DMSP into these alternative metabolite pools. Whether this type of response to sulfate deficiency is a specific feature of *E. huxleyi* or is common among diverse marine algae taxa remains to be elucidated.

ACKNOWLEDGEMENTS
M.B. was supported by a University of East Anglia (UEA) Zuckerman PhD Studentship and T.B. was supported by a Natural Environment Research Council (NERC) UK SOLAS Knowledge Transfer grant (NE/E001696/1). S.K.'s research is supported by BB/J004561/1 grant from BBSRC and the John Innes Foundation. G.M. was funded through a UK NERC Advanced Fellowship (NE/B501039/1). The RNAseq was supported by NERC Biomolecular Analysis Facility grant MGF317 and the JIC and UEA Earth and Life Systems Alliance. We are grateful to the US Department of Energy Joint Genome Institute and the scientific community who produced the E. huxleyi 1516 genome sequence. We thank Urmi Trivedi, GenePool Edinburgh, for initial bioinformatics analysis of RNAseq data. We are also grateful to Gareth Lee and Rob Utting at UEA for technical support.


expression profiling of sulfur-starved Arabidopsis by DNA macroarray reveals the role
of O-acetyl-l-serine as a general regulator of gene expression in response to sulfur
nutrition. *Plant Journal* **33**: 651-663.

Hockin NL, Mock T, Mulholland F, Kopriva S, Malin G. **2012**. The response of diatom
central carbon metabolism to nitrogen starvation is different from that of green algae and

Hoefgen R, Nikiforova VJ. **2008**. Metabolomics integrated with transcriptomics: assessing

Kopriva S, Mugford SG, Matthewman C, Koprivova A. **2009**. Plant sulfate assimilation

Koprivova A, North KA, Kopriva S. **2008**. Complex signaling network in regulation of
adenosine 5’-phosphosulfate reductase by salt stress in Arabidopsis roots. *Plant
Physiology* **146**: 1408-1420.

Lee B-R, Koprivova A, Kopriva S. **2011**. Role of HY5 in regulation of sulfate assimilation

Lyon BR, Lee PA, Bennett JM, DiTullio GR, Janech MG. **2011**. Proteomic analysis of a
sea-ice diatom: salinity acclimation provides new insight into the

Heijde M, Kaas R, La Roche J et al. **2010**. Digital expression profiling of novel
diatom transcripts provides insight into their biological functions. *Genome Biology* **11**: R85.

Maruyama-Nakashita A, Inoue E, Watanabe-Takahashi A, Yamaya T, Takahashi H.
2003. Transcriptome profiling of sulfur-responsive genes in Arabidopsis reveals global
effects of sulfur nutrition on multiple metabolic pathways. *Plant Physiology* **132**: 597-


Van Mooy BA, Rocap G, Fredricks HF, Evans CT, Devol AH. 2006. Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine


SUPPORTING INFORMATION

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

Figure S1. Fv\Fm of sulfate deficient *Emiliania huxleyi*.

Figure S2. qPCR confirmation of RNAseq data.

Figure S3. Neighbor-joining tree of putative sulfate transporters from *E. huxleyi*.

Table S1. Transcripts present in both cultures that changed expression between sulfate deficient and control *E. huxleyi* cells.

Table S2. Transcripts present only in control or in sulfate deficient cells.

Table S3. Genomic loci differentially expressed between the control and sulfate deficient *E. huxleyi* cells.

Table S4. Differentially expressed transcripts that do not align with *E. huxleyi* genome sequence.

Table S5. Functional categories enriched among *E. huxleyi* genes regulated by sulfate deficiency.

Table S6. Regulation of genes for sulfate uptake and assimilation.

Table S7. Primers used for qRT-PCR.

Supplementary Methods
Table 1. Numbers of RNAseq reads mapping to transcriptome and genome.

<table>
<thead>
<tr>
<th>Replicate</th>
<th>No. reads in dataset (after QC)</th>
<th>No. reads unmapped to transcriptome</th>
<th>% reads unmapped to transcriptome</th>
<th>No. reads unmapped to genome</th>
<th>% reads unmapped to genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>ehux25 1</td>
<td>4948310</td>
<td>1901080</td>
<td>38.42</td>
<td>933064</td>
<td>18.86</td>
</tr>
<tr>
<td>ehux25 2</td>
<td>13065126</td>
<td>4633031</td>
<td>35.46</td>
<td>2458290</td>
<td>18.82</td>
</tr>
<tr>
<td>ehux25 3</td>
<td>6978788</td>
<td>2460340</td>
<td>35.25</td>
<td>1129829</td>
<td>16.19</td>
</tr>
<tr>
<td>ehux5 1</td>
<td>14422560</td>
<td>5594499</td>
<td>38.79</td>
<td>2615982</td>
<td>18.14</td>
</tr>
<tr>
<td>ehux5 2</td>
<td>11262788</td>
<td>4208559</td>
<td>37.37</td>
<td>1945306</td>
<td>17.27</td>
</tr>
<tr>
<td>ehux5 3</td>
<td>8193958</td>
<td>3065880</td>
<td>37.42</td>
<td>1421607</td>
<td>17.35</td>
</tr>
<tr>
<td>Total</td>
<td>58871530</td>
<td>21863389</td>
<td>37.12</td>
<td>10504078</td>
<td>17.77</td>
</tr>
</tbody>
</table>
FIGURE LEGENDS

Figure 1. Growth response of Emiliania huxleyi to sulfate limitation. 
E. huxleyi cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM, 10 mM, 5 mM, and 1 mM. (a) Cell density and (b) cell volume were measured using a particle counter. (c) DMSP content was determined using gas chromatography. Results are shown as means ± standard deviation from 3 independent cultures.

Figure 2. Restoration of growth of sulfate deficient E. huxleyi by sulfate. 
E. huxleyi cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth sulfate concentration in part of the 5 mM cultures was adjusted to 25 mM by addition 20 mM sulfate. (a) Cell density and (b) cell volume were measured using a particle counter. (c) DMSP content was determined using gas chromatography. (d) APS reductase activity was determined. Results are shown as means ± standard deviation from 3 independent cultures. Different letters mark values significantly different at P<0.05.

Figure 3. Thiol content in sulfate deficient E. huxleyi. 
E. huxleyi cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. The contents of (a) cysteine and (b) glutathione were determined by HPLC and expressed per cell volume. Results are shown as means ± standard deviation from 3 independent cultures. Different letters mark values significantly different at P<0.05.

Figure 4. Sulfate uptake is induced in sulfate deficient E. huxleyi. 
E. huxleyi cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM (dark grey) or 5 mM (light grey). At a mid-log phase of growth the cells were re-suspended in transport medium, ESAW containing 25 mM or 5 mM sulfate supplemented with $^{35}$S sulfate, and cultivated for 60 min. Sulfate uptake was determined in
the cells after tissue solubilization via scintillation counting. Results are shown as means ±standard deviation from 3 independent cultures. Different letters mark values significantly different at P<0.05.

Figure 5. Summary of RNAseq analysis.

(a) The analysis pipeline used during mapping to both the *E. huxlei* transcriptome and scaffold sequences. (b) Numbers of reads and transcripts obtained during the analysis. The first three boxes refer to the combined number of reads used at each step in the analysis pipeline. The fifth box represents the numbers of transcripts assembled after mapping to the *E. huxlei* transcriptome and scaffolds and the final box is the number of differentially expressed genes identified in each of the three datasets. Round boxes represent analysis tools, square boxes show datasets.

Figure 6. Pie chart representing the KOG gene function categories of *E. huxleyi* genes up-regulated by sulfate deficiency.

*E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth the cells were harvested and gene expression was assessed by RNA sequencing. KOG functional categories were assigned to up-regulated genes (q≤ 0.05, fold change>2). Presented is the percentage of up-regulated genes assigned to individual categories from KOG annotated genes.

Figure 7. Regulation of sulfate uptake and assimilation by sulfate deficiency in *E. huxleyi*.

*E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth the cells were harvested and gene expression was assessed by RNA sequencing. Presented is the regulation of genes involved in sulfate uptake and assimilation. The results are colour-coded according to the log₂ value of ratio between transcript levels in sulfate deficient cells vs. control cells. Blue colour represents gene not present on Arabidopsis microarray. The results are compared to studies of other model organisms. Asterisk, cross and double-cross indicate transcripts from
E. huxleyi; C. reinhardtii (González-Ballester et al., 2010) and A. thaliana (Maruyama-Nakashita et al., 2006), respectively.
Figure 1. Reduction in sulfate content limits growth of *Emiliania huxleyi*. *E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM, 10 mM, 5 mM, and 1 mM. (a) Cell density and (b) cell volume were measured using a particle counter. (c) DMSP content was determined using gas chromatography. Results are shown as means ± standard deviation from 3 independent cultures.
Figure 2. Restoration of growth of sulfate deficient *E. huxleyi* by sulfate.
*E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth sulfate concentration in part of the 5 mM cultures was adjusted to 25 mM by addition 20 mM sulfate. (a) Cell density and (b) cell volume were measured using a particle counter. (c) DMSP content was determined using gas chromatography. (d) APS reductase activity was determined. Results are shown as means ± standard deviation from 3 independent cultures.
Figure 3. Thiol content in sulfate deficient *E. huxleyi*. *E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. The contents of (a) cysteine and (b) glutathione were determined by HPLC and expressed per cell volume. Results are shown as means ±standard deviation from 3 independent cultures.
Figure 4. Sulfate uptake is induced in sulfate-deficient *E. huxleyi*. *E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM (dark grey) or 5 mM (light grey). At a mid-log phase of growth the cells were re-suspended in transport medium, ESAW containing 25 mM or 5 mM sulfate supplemented with \([^{35}\text{S}]\text{sulfate}\), and cultivated for 60 min. Sulfate uptake was determined in the cells after tissue solubilization via scintillation counting. Results are shown as means ±standard deviation from 3 independent cultures.
Figure 5. Summary of RNAseq analysis.
A The analysis pipeline used during mapping to both the *E. huxlei* transcriptome and scaffold sequences. B Numbers of reads and transcripts obtained during the analysis. The first three boxes refer to the combined number of reads used at each step in the analysis pipeline. The fifth box represents the numbers of transcripts assembled after mapping to the *E. huxlei* transcriptome and scaffolds and the final box is the number of differentially expressed genes identified in each of the three datasets. Round boxes represent analysis tools, square boxes show datasets.
**Figure 6.** Pie chart representing the KOG gene function categories of *E. huxleyi* genes up-regulated by sulfate deficiency.

*E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth the cells were harvested and gene expression was assessed by RNA sequencing. KOG functional categories were assigned to significantly up-regulated genes (q ≤ 0.05, fold change>2). Presented is the percentage of up-regulated genes assigned to individual categories from KOG annotated genes.
**Figure 7.** Regulation of sulfate uptake and assimilation by sulfate deficiency in *E. huxleyi*. *E. huxleyi* cultures were grown in artificial sea water ESAW with sulfate concentrations adjusted to 25 mM or 5 mM. At a mid-log phase of growth the cells were harvested and gene expression was assessed by RNA sequencing. Presented is the regulation of genes involved in sulfate uptake and assimilation. The results are colour-coded according to the log$_2$ value of ratio between transcript levels in sulfate deficient cells vs. control cells. Blue colour represents genes not detected on a microarray. The results are compared to studies of other model organisms. Asterisk, cross and double-cross indicate transcripts from *E. huxleyi; C. reinhardtii* (González-Ballester *et al.*, 2010) and *A. thaliana* (Maruyama-Nakashita *et al.*, 2006), respectively.