Available online at www.sciencedirect.com
ScienceDirect
Gene 389 (2007) 136–145
GENE
SECTION
FUNCTIONAL GENOMICS
www.elsevier.com/locate/gene

Gene expression patterns in *Euglena gracilis*: Insights into the cellular response to environmental stress

Verónica dos Santos Ferreira [a], Iara Rocchetta [b], Visitación Conforti [b], Shellie Bench [c], Robert Feldman [c,d], Mariano J. Levin [a,e,*]

[a] *Laboratorio de Biología Molecular de la Enfermedad de Chagas (LaBMECh) INGEBI, National Research Council (CONICET), Centro de Genómica Aplicada (CeGA), University of Buenos Aires, Vuelta de Obligado 2490 2P, 1428, Buenos Aires, Argentina*
[b] *Laboratorio de Biología Comparada de Protistas Departamento de Biodiversidad y Biología Experimental–DBBE (FCEyN), Universidad de Buenos Aires, Argentina*
[c] *University of California Santa Cruz, Ocean Sciences Department, Earth and Marine Sciences Bldg. Mailroom A-316, Santa Cruz, CA 95064, USA*
[d] *SymBio Corporation, 1455 Adams Drive, Menlo Park, CA 94025, USA*
[e] *Institut Cochin, D'epartement Maladies Infectieuses, INSERM U567, Paris F-75014, France*

Received 11 July 2006; received in revised form 13 October 2006; accepted 27 October 2006
Available online 11 November 2006

Abstract

To better understand *Euglena gracilis* gene expression under different stress conditions (Chromium, Streptomycin or darkness), we undertook a survey of the *E. gracilis* transcriptome by cDNA sequencing and microarray analysis. First, we constructed a non-normalized cDNA library from the *E. gracilis* UTEX strain and sequenced a total of 1000 cDNAs. Six hundred and ten of these ESTs were similar to either Plantae or Protistae genes (*e*-value$<e^{-10}$). Second, microarrays were built by spotting all the ESTs onto mirror slides. Microarray expression analysis indicated that 90 out of those 610 ESTs changed their expression level in response to different stress treatments ($p<0.05$). In addition, we detected 10 ESTs that changed expression levels irrespective of the tested stress. These may be considered as part of a larger set of stress-related genes in *E. gracilis*. Finally, we identified 23 unknown ESTs (U-ESTs) following the expression profiles of these putative stress-related genes suggesting that they could be related to the cellular mechanism of stress response.

Keywords: cDNA library; EST; Microarray; Chromium; Streptomycin

1. Introduction

Euglena gracilis is a free-living chloroplast containing flagellated protist belonging to one of the most primitive eukaryotic groups, the euglenoids. The euglenoids include an ancient and diverse lineage of non-parasitic flagellates found predominantly in freshwater, but also occurring in marine and soil environments. They can be clearly distinguished from all other protists by a number of cytological, morphological and physiological features (Buetow, 1982). They include green phototrophic, colorless phagotrophic and osmotrophic protist species. Despite recent progress in the field of taxonomy, the phylogenetic relationships within this group are still not fully resolved (Linton et al., 2000; Müllner et al., 2001; Marin et al., 2003; Nudelman et al., 2003; Triemer et al., 2006).

On the basis of shared cytological traits, kinetoplastids, including the trypanosomatid human parasites and euglenoids are traditionally placed in a common phylum, Euglenozoa (Cavalier-Smith, 1993). This common ancestry is supported by molecular phylogenetic analysis of nuclear-encoded genes such as ribosomal RNA (Maslov et al., 1999). Some euglenoids contain a chloroplast whereas kinetoplastids do not. It has been

Abbreviations: EGM, *Euglena gracilis* medium; Sm, Streptomycin; Cr, Chromium; EST, expressed sequence tag; bp, base pairs; Cy, cyanine dye; SDS, sodium dodecyl sulphate; SSC, sodium chloride sodium citrate; U-ESTs, unknown ESTs.

* Corresponding author. Instituto de Investigaciones en Ingeniería Genética y Biología Molecular (INGEBI), Vuelta de Obligado 2490, 1428, Buenos Aires, Argentina. Tel.: +54 11 4783 2871; fax: +54 11 4786 8578.
E-mail address: mlevin@dna.uba.ar (M.J. Levin).

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.10.023

argued that chloroplast arose early in the diversification of the lineage Euglenozoa. These views are collectively referred as the plastid early hypothesis (Gibbs, 1978; Margulis, 1993; Leander, 2004). There are some molecular evidences that kinetoplasts possessed a plastid at some point of their evolutionary history but they subsequently lost it retaining numerous plastid genes (Hannaert et al., 2003). *E. gracilis* possesses chloroplasts and is presumed to have a mosaic genome, although little is known about its genome size and structure (Russell et al., 2005).

E. gracilis cells are known to be sensitive to chemical and physical factors, such as Chromium (Cr), Streptomycin (Sm) and the absence of light, which can inhibit the development of chloroplast and the synthesis of chlorophyll (Gajdosova and Reichrtova, 1996).

Chromium is a highly toxic non-essential metal for microorganisms and plants. Due to its widespread industrial use, Cr has become a pollutant in diverse environments settings (Cervantes et al., 2001). The presence of Cr in the environment may pose a strong selective pressure on a microorganism's ability to tolerate this contaminant. Chromium toxicity appears to be related to the production of Reactive Oxygen Species (ROS), a phenomenon known as ROS effect (McCord and Fridovich, 1969). The mechanism underlying different levels of Cr sensitivity among different algal taxa remains to be elucidated (Rocchetta et al., 2003).

Streptomycin is an aminoglycoside antibiotic with a known bleaching action on *E. gracilis* (Schwartzbach and Schiff, 1974). It brings about the permanent loss of plastids and plastid DNA in dividing photosynthetic cells and blocks the development of the chloroplast in non-dividing cells, but has no effect on cell division or viability. This effect may be linked to the binding of Sm to the chloroplast ribosomes, resulting in a selective inhibition of plastid protein synthesis (Schwartzbach and Schiff, 1974).

Interestingly, *E. gracilis* cells grown under constant darkness conditions also lose their chloroplast but in a reversible manner (Schiff and Schwartzbach, 1982). The molecular mechanisms underlying the adaptative response to the presence of Cr, Sm and the absence of light are unknown.

To test the feasibility of a genomic approach to study the algal sensitivity to Cr, Sm and the absence of light, we designed a small scale strategy of cDNA sequencing and microarray-based expression analysis that would allow us to identify *E. gracilis* genes that are differentially expressed under different environmental conditions. The results presented herein have to be considered as a first step in the understanding of gene function, physiological control and developmental processes in *E. gracilis*.

2. Materials and methods

2.1. Culture conditions and bioassays

E. gracilis commercial strain (UTEX 753) came from the Culture Collection of Algae of Texas University, USA (generously provided by Dr Richard Triemer) and *E. gracilis* wild type strain (MAT) was isolated from Matanza River by Lic. Laura Ruiz from the Laboratory of Comparative Biology of Protists (Ruiz et al., 2004). Cells from both strains were grown on organic EGM medium (sodium acetate 1 g, beef extract 1 g, tryptone 2 g, yeast extract 2 g, calcium chloride × 2 H_2O 0.01 g, dissolved in 1 l of distilled water) in a growth chamber at 24 ± 1 °C, 16/8 h L/D (light/dark). Each novel unialgae culture was initiated 6 days before experiments in order to obtain an inoculum in exponential growth. Axenicity was monitored plating *E. gracilis* cultures in Bacto Brain Heart infusion (FisherBiotech) to control for bacterial or yeast contaminations.

All experiments were carried out under the same growing conditions, on static cultures containing 150 ml culture medium in 250-ml glass flasks, at 24 ± 1 °C, 16/8. Aliquots of 10^5 cells ml^{-1} from both stock cultures, MAT and UTEX, were inoculated for Cr treatment into each glass flask. $K_2Cr_2O_7$ was added axenically from a stock solution until reaching total dichromate concentrations of 100 μM Cr (VI). Assays lasted 96 h (U.S. Environmental Protection Agency 1985). *E. gracilis* was bleached with streptomycin (100 μg/ml, 7 days) (SmMAT) and absence of light (1 month, dMAT).

2.2. Isolation of RNA and cDNA synthesis

Total RNA was prepared from saturated untreated UTEX cultures, using a hot phenol method (Aiba, 1985). Total RNA integrity was assayed by 1.2% agarose/formaldehyde gel electrophoresis as described by Greenwood and Gray (1998). Total RNA concentration was determined by measurements at 260 nm and 280 nm. The Poly $(A)^+$ Pure-Ambion kit was used to prepare mRNA from total RNA. For RT reactions and PCR amplification, the following primers were alternatively used: i) RT: random decamers/PCR: random decamers; ii) RT: oligo (dT) anchor as designed by Ben-Dov et al. (2005)/PCR: oligo (dT) anchor and SL or miniexon (Frantz et al., 2000) from *E. gracilis*. The PCR program was: 94 °C 4 min (1 cycle), 94 °C 50 s, 50 °C 1 min, 72 °C 1 min (32 cycles), 72 °C 10 min final extension.

2.3. cDNA library. ESTs sequencing

The cDNA products were subcloned in pGEM-Teasy Vector System I (Promega). The library was used to transform DH10B *E. coli* by electroporation, plated on ampicillin agar Petri dishes, and colonies were picked for sequencing. All picked colonies were grown overnight in 96-well plates for no more than 12 h before being directly archived at −80 °C. PCR reactions used M13 forward and reverse primers, and following shrimp alkaline phosphatase/exonuclease I treatment, products were directly sequenced with T7 or Sp6 primer on MegaBace 1000 and 500 sequencer. Archived clones are available upon request.

2.4. Bioinformatics

Raw sequence trace data were processed to screen out vector and linking sequence, to remove low-quality sequence, and to trim poly(dA) tails using an in-house software solution. Phred was used to assess quality of reads longer than 200 bp (Ewing et al.,

1998). The sequences were annotated according to similarity to identify genes in different databases and submitted to dbEST. To identify the non-redundant set of putative genes, sequences were clustered on the basis of sequence similarity using the CLOBB program (Parkinson et al., 2002). BLAST analyses: BLASTN against a non-redundant DNA database (GenBank); BLASTX against a non-redundant protein database (SwissProt-trEMBL) and BLASTN against dbEST. Results from these analyses are available upon request. Similarity between sequences with an *e*-value $< e^{-10}$ was considered as significant.

2.5. Construction of cDNA microarrays

cDNA was derived from untreated UTEX *E. gracilis*. The cDNAs were PCR amplified and purified using Corning FilterEXTM 384 well filter plates (Corning, USA). The quality and quantity of purified PCR products were confirmed using agarose gel electrophoresis. More than 90% of the clones gave a single band. Purified PCR products were dried and resuspended in 10 μl of spotting solution (Microarray Crosslinking Reagent D Amersham Biosciences, USA, GE). cDNAs solutions were spotted onto silianized microarrays slides (type 7 Star, Amersham Biosciences, GE) using a capillary pen type arrayer (Generation III, Amersham Biosciences, GE). cDNA spotted slides were then exposed to 50 mJ of 250 nm light to crosslink DNA on slides. Each slide was spotted with duplicate sets of unique cDNAs spotted on the left- and right-hand sides of the array, in total 2000 cDNAs were spotted on 24 slides in duplicate. Lambda phage DNA (Takara, Japan) was used as an external control gene and was spotted onto each array block (32 × 12) with more than one spot. Spotting solution without DNA was spotted as a negative control on each block.

2.6. Preparation of probes and hybridization to cDNA microarray

Fluorescently labeled probes were generated by reverse transcription as follows: 10 μg of total RNA from untreated cells, 2.5 ng of green-fluorescent protein poly(A)RNA, 8 pmol of anchored dT primer, and 1 μg of random nonamers were combined in a 10–15 μl reaction volume. The solution was heated to 70 °C for 10 min, chilled briefly on ice, and centrifuged. Reverse transcription was performed in a 20 μl reaction volume. Final concentration were 1× first strand buffer (Life Technologies, Rockville, Md.), 10 mM dithiothreitol, 100 nM dATP, dGTP and dTTP, 50 nM non-labeled dCTP, 50 nM FluoroLink-dCTP (either Cy3 or Cy5 labeled, Amersham CyScribe First-Strand cDNA Labeling Kit with CyScribe GFX Purification Kit®), and 0.5 U/μl of placental RNase inhibitor (Promega, Madison, Wis.). The contents were mixed and incubated at room temperature for 10 min. Superscript II RT (Life Technologies) was added (200 U), and the reaction mixtures were incubated at 42 °C for 2 h. RNA was hydrolyzed with sodium hydroxide (0.25 N final concentration) for 15 min at 37 °C. Samples were neutralized by the addition of 2 M MOPS buffer to a final concentration of 0.4 M. Probes were concentrated using a vacuum centrifuge dryer. The labeling reaction efficiency of each cDNA probe was confirmed by spectrophotometry, measuring the absorbance at 260 550 and 650 nm. Prior to hybridization, microarray slides were pretreated for 20 min at 55 °C in 5× SSC −0.2% SDS, rinsed briefly in deionized water, and dried with compressed air. Hybridization reactions were performed using a manufacturer's hybridization buffer (Amersham Biosciences-GE) containing formamide for 16 h at 55 °C in an air phase incubator. After hybridization, the slides were washed in 2× SSC/0.1% SDS for 10 min, followed by washing in 1× SSC/0.1% SDS for 5 min each at 55 °C, then in 0.1× SSC at room temperature to remove SDS. Spinning at low speed in a centrifuge for 1 min dried the slides.

2.7. Scanning and data processing

The slides were scanned using a GenePix4000B DNA microarray scanner (Axon instrument, USA) at ~635 nm (Cy5) and then at ~532 nm (Cy3). The voltage of both channels of the photo multiplayer tubes was adjusted to obtain a ratio value of 1.0 for the external control spots. The resulting fluorescent intensities for each spot were quantified by an algorithm for adaptive circle segmentation of an image, and signal intensities calculated from the median of the background intensities subtracted from the median of the intensities of each pixel ((F532 Median) − (B532) or (F635 Median) − (B635)) using GenePix Pro 4.0 microarray analysis software (Axon Instruments). Two independent hybridization experiments, each using one set of slides, were performed. An intensity-dependent normalization (non-linear of LOWESS normalization) method (Clevel and Devlin, 1988) was applied to normalize the data from replicate experiments using Genespring software version 5.0 (Sillicon Genetics, USA). The fluorescence units of each duplicate spot on each slide were averaged and the ratio value was then calculated; so that the signal intensity of each tested sample was divided by that of the control sample. We labeled the same RNA with different dyes to evaluate the variation of the results depending on the labeled dye bias, and also to determine the cutoff values for up- and downregulated genes. Under our experimental conditions, we set the significance value at >1.8 for upregulated and <0.5 for downregulated genes, respectively.

2.8. Northern blot analysis

Northern blots were performed as described previously (Agy et al., 1990). Radioactive probes for northern analysis were generated from random selected clones. Hybridization was carried out at 68 °C for 16 h using ^{32}P-labeled probes (RediPrime Kit, Amersham-GE). Blots were washed 3–4 times in 2× SSC, 0.1% SDS at 50 °C and exposed to a phosphorstorage screen (Molecular Dynamics®) without drying. Computer generated images (MD Phosphor-Imager 400S) of individual gels were analyzed using ImageQuant Software (Kucharski and Maleszka, 2002).

3. Results and discussion

3.1. Construction of a cDNA library from E. gracilis UTEX 753. cDNA sequence analysis and gene annotation

A non-normalized cDNA library was generated from the *E. gracilis* commercial strain UTEX 753. During the library construction, size fractionation was used to enrich in full-length

Table 1
Summary of *E. gracilis* EST analysis

	Number of ESTs	%
Total ESTs	1000	100
Non-redundant sequences	910	91
Redundant sequences	90	9
Putative identified genes	610	61
Putative unknown genes	39	39
Average size of ESTs	500 pb	
Functional classification of known genes		
Chloroplast proteins	240	24
Transcription/translation initiator factors	40	4
Structural proteins	100	10
Transport proteins	80	8
General metabolism	110	11
Mitochondrial proteins	20	2
Ribosomal proteins	20	2
Total	610	61

cDNAs. One thousand cDNA fragments were cloned and bidirectionally sequenced. Sequence length ranged from 100 to 700 pb, with an average length of 500 bp.

To identify their potential biological function, sequences were translated into their six possible reading frames and compared to the GenBank non-redundant protein database using BLASTX. Six hundred and ten ESTs showed similarities to proteins with known function (61%, *e*-value $< e^{-10}$) and 390 (39%) showed no similarity to any other known polypeptide (Table 1). The total sequenced ESTs showed only 9% redundancy (see Section 2.4). These redundant sequences were transcripts of the same gene or cognate genes (Table 1).

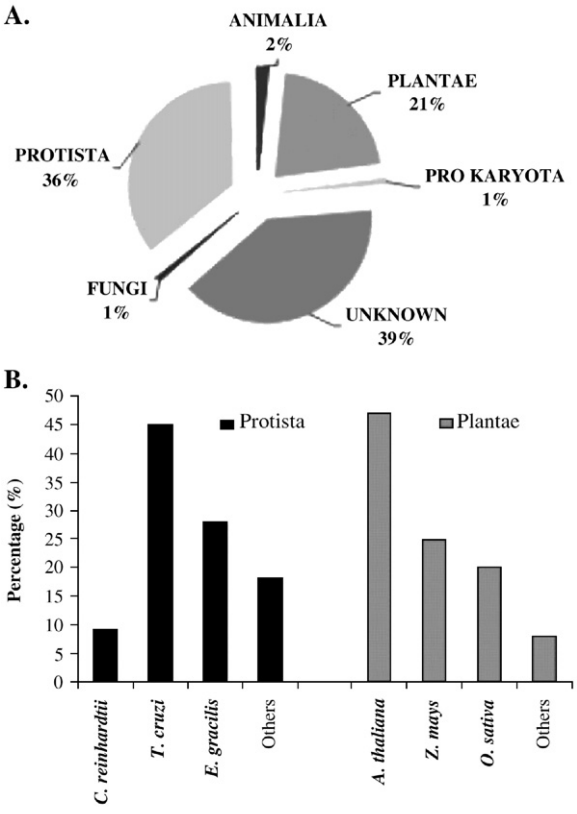

Fig. 1. *E. gracilis* ESTs BLAST search: A. database match taxonomic categories of ESTs sequenced in *E. gracilis* and B. database match of *E. gracilis*' ESTs within different species: *Chlamydomonas reinhardtii*, *Trypanosoma cruzi*, *Euglena gracilis*, *Arabidopsis thaliana*, *Zea mays* and *Oriza sativa*.

3.1.1. Functional classification of E. gracilis ESTs

The 610 distinct ESTs similar to proteins with known function were organized in seven major categories, as follows (Table 1): chloroplast proteins, structural proteins, transport proteins, proteins involved in general metabolism, transcription/translation factors, mitochondrial and ribosomal proteins. The complete list of sequences was deposited in the NCBI EST database (www.ncbi.nlm.nih.gov/projects/dbEST) and is available with the following accession numbers: DN976397–DN976420, EC591293–EC591315 and EC609911–EC611866.

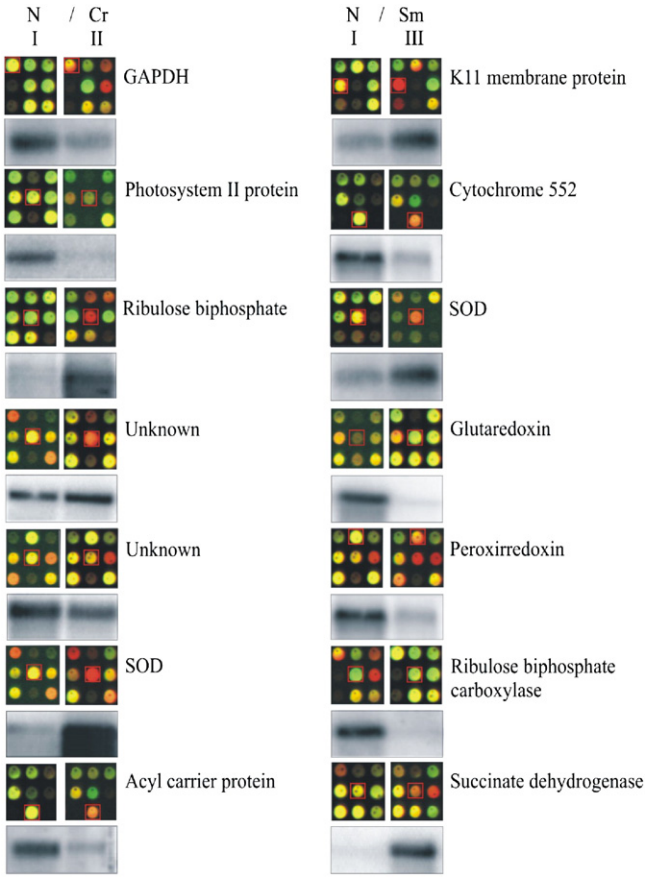

Fig. 2. Differential expression of fourteen (14) ESTs. Two columns on the left correspond to transcripts identified as differentially expressed after Chromium treatment (N I: normal treatment, Cr II: Chromium treatment). The two columns on the right represent expression after Streptomycin treatment (N I: normal treatment, Sm III: Streptomycin treatment). Each EST is identified by its putative function. The expression changes of the spot representing the expression of the corresponding EST are marked by a red square. Below the spots, northern blots indicate expression of the mRNA recognized y the EST probe. The northern blots show hybridization of ESTs to membranes with (I) *E. gracilis* RNA from cells grown in normal conditions, (II) *E. gracilis* RNA from cells grown in the presence of Chromium and (III) *E. gracilis* RNA from cells grown in the presence of Streptomycin. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The total hits found for *E. gracilis* ESTs within different taxonomic kingdoms are shown in Fig. 1A. Within the group of 610 ESTs similar to known proteins (known ESTs), we identified and focused on two main sub-groups: those ESTs related to plants and unicellular algae (Plantae-algae group, 43%, 260 out of 610 known ESTs) and those related to kinetoplastids (Kinetoplastidae group, 34%, 210 out of 610 known ESTs). This classification is consistent with *E. gracilis*' common nuclear lineage with kinetoplastids and a chloroplast lineage, derived from green algae (Yasuhira and Simpson, 1997; Nozaki et al., 2003). Fig. 1B shows species having the highest number of matches to *E. gracilis* ESTs.

3.2. Construction of microarrays and confirmation of array results

After characterizing ESTs, microarrays were built by spotting them onto mirror slides. A first set of experiments was designed to assess the correlation of data originated using the microarrays with the expression of genes as determined by northern blots (Ramdas et al., 2001).

Previous assays demonstrated that the UTEX strain did not withstand treatments representing putative environmental stress conditions. However, MAT was found to be very resistant to environmental stimuli (Rocchetta et al., 2003). Due to this, microarrays of UTEX ESTs, were probed with cDNAs derived from the stress resistant strain MAT *E. gracilis* grown normally and under three different stress conditions (CrMAT, SmMAT and dMAT, Section 2.1).

Cye-labeled probes generated from *E. gracilis*' total RNA grown in normal conditions and under different stimuli (CrMAT, SmMAT and dMAT) were hybridized to the microarrays. Pair wise comparison of results obtained using probes from cells grown with and without stimulation allowed identification of a first set of differentially expressed ESTs.

To confirm their differential expression, we selected 14 EST that showed notable variation in their expression level, seven after Cr and seven after Sm treatment. Northern blots and the corresponding microarray spots are shown in Fig. 2. After Cr treatment, the increased expression of the mRNAs for ribulose biphosphate carboxylase (DN976410), superoxide dismutase (SOD) (DN976412) and for two genes with unknown function (EC611865, EC611825) was confirmed by northern blots. The decreased expression of genes encoding for glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (DN976404), a protein of the photosystem II (DN976409) and an acyl carrier protein (DN976411) were also confirmed by northern blots (Fig. 2). As a consequence of the Sm treatment, the mRNA levels of membrane protein K11 (DN976413), succinate dehydrogenase (DN976416) and SOD (DN976412) increased, whereas there was an underexpression for ribulose biphosphate carboxylase (DN976410), cytochrome 552 (DN976407), glutarredoxin (DN976418) and peroxirredoxin (DN976417) (Fig. 2). These results validated our microarray approach and prompted a first analysis of *E. gracilis* gene expression.

3.3. Identification of changes in gene expression using microarrays

To determine the basal level of expression for both UTEX and MAT *E. gracilis* strains, microarray I was hybridized to a Cy3-UTEX and a Cy5-MAT probe from cultures grown at 25 °C, without any type of stress ("reference" or normal slide, Fig. 3). Whereas, results from microarrays II, III and IV were generated with the following no-stress/stress probe combinations: Cy3-MAT/Cy5-CrMAT, Cy3-MAT/Cy5-SmMAT and Cy3-MAT/Cy5-dMAT. All the data generated from microarrays II, III and IV were compared to the basal expression levels obtained from microarray I. Interestingly, when *E. gracilis* cells were grown without stress, the expression levels of the genes included in this array were similar for UTEX and MAT.

Data from 24 independent co-hybridizations were analyzed to identify ESTs presenting differences in gene expression. For each probe combination, the experiment was repeated reversing the fluorescent dye. Thereafter, we calculated the ratios of fluorescence intensities of the two probes and determined the number of clones that were differentially expressed showing a ratio greater than 1.25-fold and a p-value <0.05. The relationship of the average hybridization intensity for each stress treatment was represented with different scatter plots

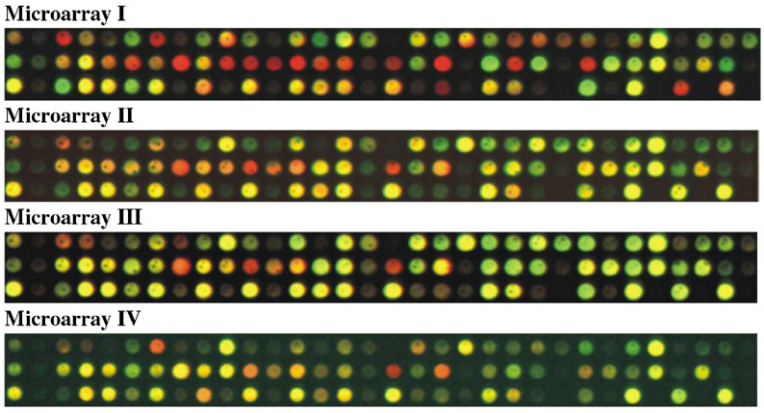

Fig. 3. Partial view of the whole array (ninety-six spots) for each microarray signal generated after hybridization with the following combination of probes: Cy3-UTEX/Cy5-MAT (microarray I), reference, Cy3-MAT/Cy5-CrMAT (microarray II), Cy3-MAT/Cy5-SmMAT (microarray III) and Cy3-MAT/Cy5-dMAT (microarray IV).

generated by the expression analysis software (see Section 2.7) (data not shown).

3.3.1. ROS response induced by Cr treatment

Fig. 4 shows the changes in expression level of *E. gracilis* ESTs within the Plantae-algae and Kinetoplastidae groups (Fig. 4A and B). The adaptation to Cr caused an increase in the expression level of the following *E. gracilis* mRNAs: SOD (DN976412), peroxirredoxin (DN976417), glutaredoxin (DN976418), copper-like chaperonine (DN976401), ribulose diphosphate carboxylase (DN976410), digalactosyl diglyceride synthetase (DN976403), cytochrome 552 (DN976407), succinate dehydrogenase (DN976416), cyclophilin (DN976415), membrane protein K11 (DN976413) and ribosomal proteins P1 (DN976405) and P2 (DN976406) and a decrease in the mRNA level of ATPase (DN976419), GAPDH (DN976404), photosystem II protein (DN976409) and acyl carrier protein (DN976411).

The most relevant consequence of the oxidative state provoked by Cr treatment was the overexpression of SOD gene. The product of this gene is one of the main natural antioxidant enzymes, and is found in cyanobacteria, higher plants and most algae (Bowler and Chua, 1994). In our case, *E. gracilis* presented two different SOD genes, one homologous to plant SOD (FeSOD) (DN976412) and another one homologous to *T. cruzi* SOD (FeSOD) (EC611866). Both SOD genes had an increased transcript level.

In addition, Cr treatment induced an overexpression of peroxiredoxin and glutaredoxin, two proteins involved in detoxification of cell peroxides and repair of cell membrane

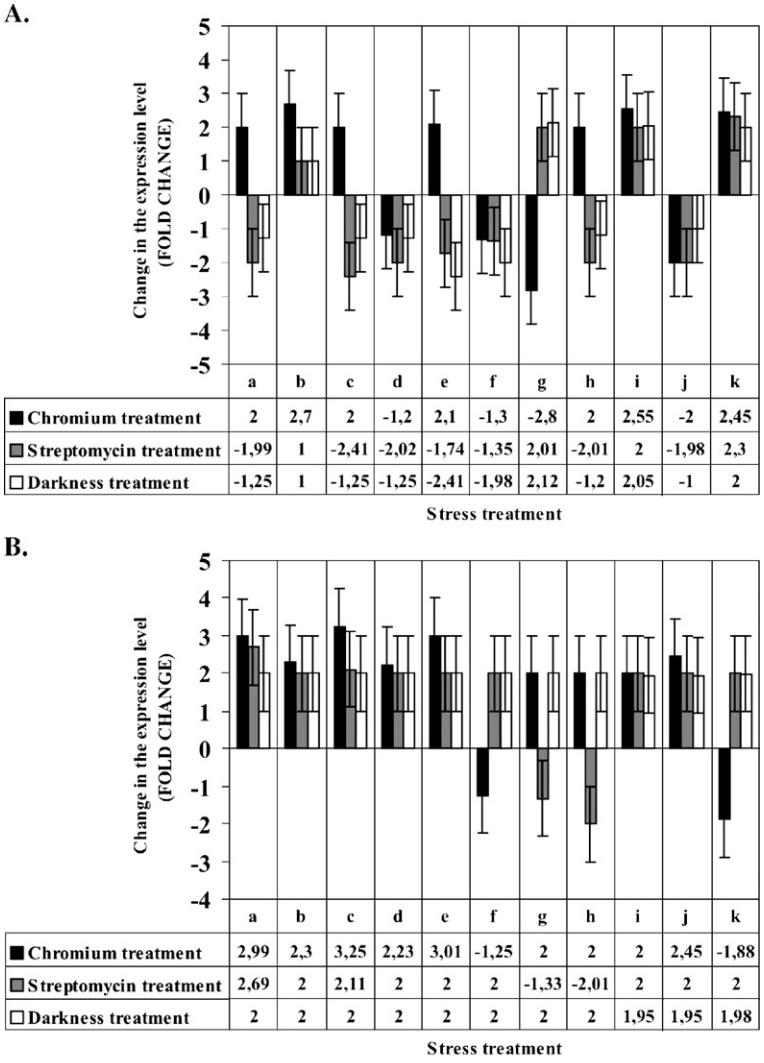

Fig. 4. Changes in the expression level of selected ESTs are shown as fold increase or decrease for each treatment. A. Analysis of expression of genes related to Plantae-algae group. Effect of Chromium, Streptomycin and darkness on eleven *E. gracilis* ESTs: a. ribulose biphosphate carboxylase, b. SOD, c. cytochrome 552, d. photosystem II protein, e. digalactosyl diglyceride synthetase, f. acyl carrier protein, g. GAPDH, h. chloroplast RF6 protein, i. copper-like chaperonine, j. chloroplast protein, k. P1 ribosomal protein; B. microarray analysis within spots with hits in the Kinetoplastidae group. The effect of environmental stress condition (Streptomycin, Chromium and darkness) on 11 *E. gracilis* EST is shown. a. K11 membrane protein, b. KMP10 membrane protein, c. SOD, d. cyclophilin, e. succinate dehydrogenase, f. GAPDH, g. peroxiredoxin, h. glutaredoxin, i. P2 ribosomal protein, j. P1 ribosomal protein, k. H-ATPase. Relative fold increase or decrease (−) is shown in bottom lines.

function after oxidative stress, respectively (Donahue et al., 1997). Similarly, we identified an increase of the mRNA level of a copper chaperonin, that may also be considered as a marker for oxidative stress (Cabiscol et al., 2002). Morphological studies revealed that Cr seems to cause a change in chloroplast shape and size (Rocchetta et al., 2003). Accordingly, two chloroplast proteins and a type of galactolipid synthetase showed increased mRNA expression: ribulose diphosphate carboxylase, the enzyme that is responsible for carbon dioxide fixation in photosynthesis, cytochrome 552, a protein marker of chloroplast protein synthesis, and digalactosyl diglyceride synthetase, an enzyme playing a role in galactolipids synthesis for the stabilization of chlorophyll-containing photoreceptor membranes (Rosenberg, 1976). Other overexpressed genes were succinate dehydrogenase, a mitochondrial marker enzyme for the electron transport pathway, cyclophilin, a protein involved in protein trafficking and maturation (Romano et al., 2004) and membrane protein K11, that may be related to the Cr-induced morphological changes of *E. gracilis* cells (Fig. 4).

Interestingly, we found increased mRNAs levels for two ribosomal P proteins, P1 and P2. As suggested by Schawalder et al. (2004), these ribosomal proteins may be involved in the stress response.

While, the decreased expression of the H^+-ATPase mRNA may be explained as a response to Cr-induced membrane depolarization (Manusadzianas et al., 2002), the decreased mRNA level for *E. gracilis* GAPDH genes, one 75% similar to Plantae-algae GAPDH (chloroplast located) (DN976408) and the other 60% identical to the Kinetoplastidae GAPDH (glycosome located) (DN976404) is presumably a consequence of alterations of the glycolytic function induced by the metal (Henze et al., 1995).

3.3.2. Streptomycin treatment: cell response to irreversible bleaching

Gene expression analysis after Sm treatment showed that the two *E. gracilis* GAPDH genes (DN976404, DN976408) were upregulated probably caused by autotrophic metabolism arrest. These results, together with the overexpression of a ATPase subunit (DN976420), agreed with the increase in the rate of glycolysis (Guerrieri et al., 1984). As expected, we identified a downregulation of several bleaching-related genes: ribulose diphosphate carboxylase (DN976410), cytochrome 552 (DN976407), digalactosyl diglyceride synthetase (DN976403), and apoproteins (DN976399, DN976400) and two photosystem precursor proteins (DN976397, DN976398, Fig. 4A and B).

3.3.3. Darkness treatment or reversible bleaching

The results obtained from exposure of *E. gracilis* cells to darkness were similar to those from Sm treatment. However, the changes in mRNA expression were not as marked as those from irreversible bleaching (Fig. 4A and B).

3.3.4. Stress-related ESTs

Gene expression analysis from the three stress conditions tested (Cr, Sm and darkness) allowed the identification of 10 genes that changed their expression pattern irrespective of the stress treatment and therefore seemed to respond to a broader spectrum of different kinds of stress. Eight of them were upregulated including SOD, membrane protein K11 and K10, cyclophilin, chaperonin, succinate dehydrogenase and ribosomal proteins P1 and P2, while acyl carrier protein and a photosystem II protein were downregulated.

3.3.5. Unknown ESTs (U-ESTs): correlation with the expression of ESTs with putative functions

Clear changes in mRNA expression corresponding to ESTs with unknown function were detected for Cr and Sm treatments (Fig. 5A). Interestingly, 23 U-ESTs presented the same expression profile with both treatments. Seventeen U-ESTs increased and six U-ESTs decreased their mRNA levels. When we compared their expression pattern with those ESTs with known function, they could be classified into five sub-groups. Eight U-ESTs followed the expression of SOD (EC611865, EC611844, EC611839, EC611837, EC611825, EC611768,

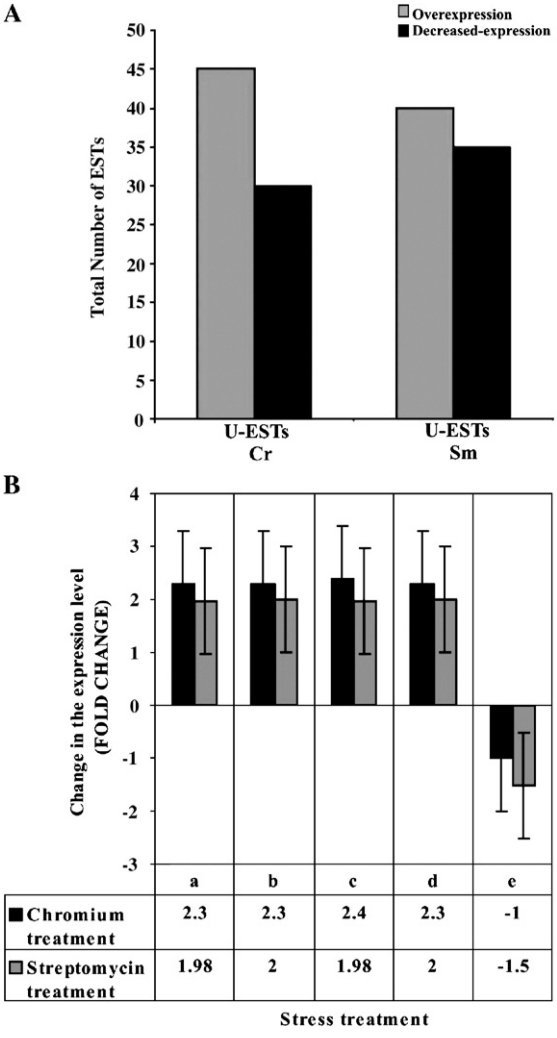

Fig. 5. A. Number of ESTs with unknown identity showing changes in expression level after the chromium and streptomycin treatments; B. putative classification of ESTs of unknown identity according to the expression pattern of ESTs with putative function (see Section 3.3.5): a. SOD, b. chaperonin, c. K11 membrane proteins, e. succinate dehydrogenase, f. acyl carrier proteins. Relative fold increase or decrease (−) is shown in bottom lines.

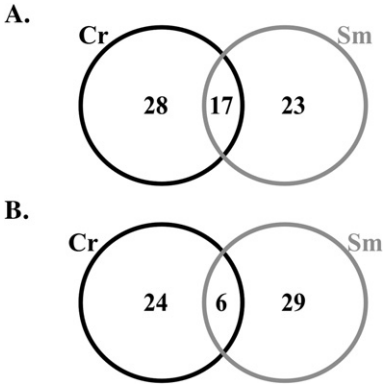

Fig. 6. Expression of mRNAs corresponding to U-ESTs that vary in the same way irrespective of the treatment. A. Increased expression after Chromium (Cr) or Streptomycin (Sm) treatment; B. decreased expression after Chromium (Cr) or Streptomycin (Sm) treatment.

EC611762, EC611756), three that of chaperonin (EC611718, EC611710, EC611789), four that of membrane protein K11 (EC611626, EC609933, EC609935, EC609941), two that of succinate dehydrogenase (EC609952, EC609954) and six followed the variation of acyl carrier proteins (EC609971, EC609974, EC610016, EC610025, EC610031, EC610053) (Fig. 5B). Because these U-ESTs followed the expression of genes reacting to stress conditions, it is plausible that they may be involved in the cellular mechanisms of stress response. In Fig. 6, these U-ESTs are represented by the shared portions of the circles.

In addition, Cr induced the increase of expression of 28 U-ESTs, whereas Sm provoked increased expression of 23 different U-ESTs (Fig. 6A). On the other hand, Cr treatment induced a decrease in the expression of 24 ESTs, and Sm that of 29 different U-ESTs (Fig. 6B).

3.4. Conclusions

After different treatments, 60 out of the 350 *E. gracilis*' plant-related ESTs showed significant changes in their expressions ratios. On the other hand, 30 out of 160 ESTs, with similarity to kinetoplastid genes showed changes in their expression ratios. Fig. 7A shows the proportion of *E. gracilis* ESTs showing changes after stress treatment. The number of spots identified after each treatment for Plantae-algae and Kinetoplastidae related genes is shown in Fig. 7B and C.

Using microarray technology for the analysis of different environmental stress conditions, we detected 10 ESTs that changed their expression levels irrespective of the type of treatment. These ESTs could be considered as part of a larger set of stress-related genes. Among these, eight genes including SOD, membrane protein K11 and K10, cyclophilin, chaperonin, succinate dehydrogenase and ribosomal proteins P1 and P2 increased their mRNA levels upon stress while acyl carrier protein and a photosystem II protein reacted by decreasing the corresponding mRNA level.

The SOD gene was expected to show overexpression as a result of oxidative stress, as it has been demonstrated in Ernani et al. (2003). Accordingly, *E. gracilis* SOD mRNA reached its highest overexpression after heavy metal treatment, the most extreme of the conditions tested.

The different types of SOD genes, as well as those encoding GAPDH proteins, support the hypothesis of a mosaic origin for *E. gracilis* genome (Henze et al., 1995).

The expression level for membrane proteins K11 and K10-like mRNAs was associated with changes of cell morphology after each treatment.

Overexpression of genes encoding for proteins that play a role in protein folding like the cyclophilin and chaperonins, was evident mainly after Cr treatment, in association with an

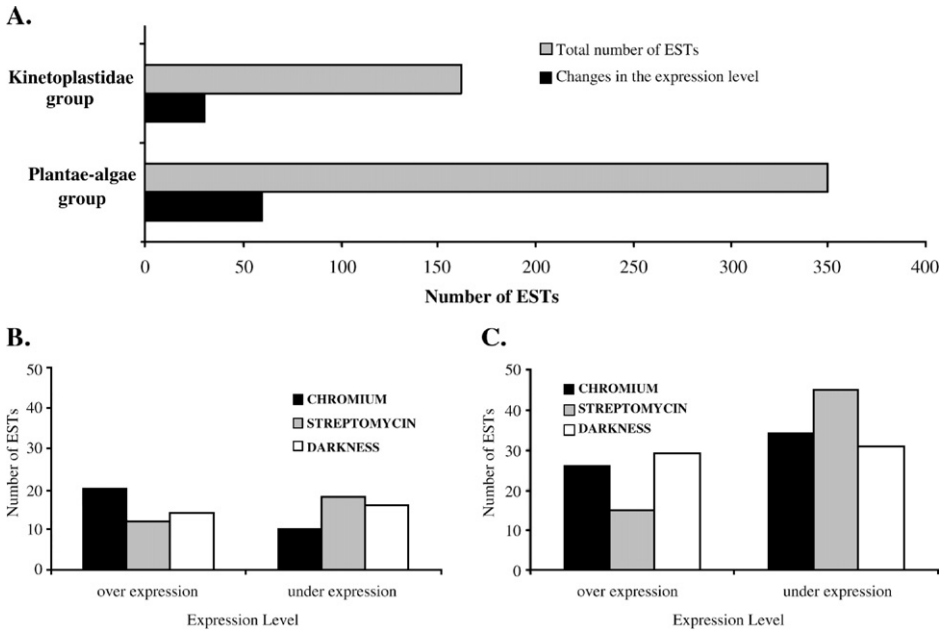

Fig. 7. A. Total number of *E. gracilis* ESTs within the selected groups and changes in their expression after the stress treatments. Number of ESTs showing changes in the level of expression after each treatment: B. within ESTs homologous to Plantae-algae group and C. ESTs homologous to Kinetoplastidae group.

increase in general metabolism and the overexpression of succinate dehydrogenase mRNA. In contrast, acyl carrier mRNA decreased after each treatment, in agreement with the apparent lack of fatty acid synthesis under stress (Ohlrogge, 1997).

Our analysis allowed the characterization of 23 unknown genes that vary their expression irrespective of the stress condition, suggesting that their gene products may be a part of the set of genes involved in the response to environmental stress, i.e. stress-related genes.

With the cloning of *E. gracilis* cDNAs, we have generated a novel tool to improve the molecular characterization of several biochemical and biological processes in which this microalgae is involved, as well as its mosaic genome.

Acknowledgements

We thank Mr. Wilson Lew for comments that greatly helped the microarray interpretation and Mr. A. Ghadiri, SymBio Corporation, for the help given during the sequencing of ESTs. We are also indebted to Dr. Hernán Lorenzi, The Institute for Genomic Research, for helpful discussions and critical reading of the manuscript.

For the period 2005–2006 MJL is International Professor of a Chaire Internationale de Recherche Blaise Pascal, Fondation Ecole Normale Superieure, Region Ile de France, Paris, France. This work was also supported by (i) World Health Organization, Special Program for Research in Tropical Diseases (TDR), South-South Initiative; (ii) FONCYT-BID 1201/OC-AR 05-6802, SECyT-Buenos Aires, Argentina; (iii) UBACyT grant X-624; (iv) PME n°137, SECyT-Buenos Aires, Argentina. The work of MJL was partially supported by an International Research Scholar grant from the Howard Hughes Medical Institute, Chevy Chase, MD, USA.

References

Agy, M.B., Wambach, M., Foy, K., Katze, M.G., 1990. Expression of cellular genes in CD4 positive lymphoid cells infected by the human immunodeficiency virus, HIV-1: evidence for a host protein synthesis shut-off induced by cellular mRNA degradation. Virology 177 (1), 251–258.

Aiba, H., 1985. Transcription of the *Escherichia coli* adenylate cyclase gene is negatively regulated by cAMP receptor protein. J. Biol. Chem. 260, 3063–3070.

Ben-Dov, C.P., Levin, M.J., Vazquez, M.P., 2005. Anlysis of the highly efficient pre-mRNA processing region HX1 of *Tripanosoma cruzi*. Mol. Biochem. Parasitol. 140 (1), 97–105.

Bowler, C., Chua, N., 1994. Emerging themes of plant signal transduction. The Plant Cell 6, 1529–1541.

Buetow, D., 1982. The Biology of *Euglena*, vol. III. Academic Press, New York, pp. 200–201.

Cabiscol, E., Belli, G., Tamarit, J., Echave, P., Herrero, E., Ros, J., 2002. Mitochondrial Hsp60, resistance to oxidative stress, and the labile iron pool are closely connected in *Saccharomyces cerevisiae*. J. Biol. Chem. 277 (46), 44531–44538.

Cavalier-Smith, T., 1993. Kingdom Protozoa and its 18 phyla. Microbiol. Rev. 57 (4), 953–994.

Cervantes, C., et al., 2001. Interactions of Chromium with microorganisms and plants. FEMS Microbiol. Rev. 25 (3), 335–347.

Clevel, W.S., Devlin, S.J., 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. 83, 596–610.

Donahue, J.L., Okpodu, C.M., Cramer, C.L., Grabau, E.A., Alscher, R.G., 1997. Responses of antioxidants to paraquat in pea leaves (relationships to resistance). Plant Physiol. 113 (1), 249–257.

Ernani, P., Sigaud-Kutner, T., Leitao, M., Okawoto, O., Morse, D., Colepicolo, P., 2003. Heavy metal induced oxidative stress in algae. J. Phycol. 39, 1008–1018.

Ewing, B., Hillier, M., Wendl, M., Green, P., 1998. Base-calling of automated sequencer traces using Phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Frantz, C., Ebel, C., Paulus, F., Imbault, P., 2000. Characterization of trans-splicing in Euglenoids. Curr. Genet. 37 (6), 349–355.

Gajdosova, J., Reichrtova, E., 1996. Different growth response of *Euglena gracilis* to Hg, Cd, Cr and Ni compounds. Anal. Bioanal. Chem. 354 (5–6), 641–642.

Gibbs, S.P., 1978. The chloroplast of *Euglena* may have evolved from symbiotic green algae. Can. J. Bot. 56, 2883–2889.

Greenwood, S., Gray, M., 1998. Processing of precursor rRNA in *Euglena gracilis*: identification of intermediates in the pathway to a highly fragmented large subunit rRNA. Biochim. Biophys. Acta 1443, 128–138.

Guerrieri, F., Micelli, S., Massagli, C., Gallucci, E., Papa, S., 1984. Interaction of the aminoglycoside antibiotic dihydrostreptomycin with the H+-ATPase of mitochondria. Biochem. Pharmacol. 33 (15), 2505–2510.

Hannaert, V., et al., 2003. Plant-like traits associated with metabolism of *Trypanosoma* parasite. Proc. Natl. Acad. Sci. U. S. A. 100 (3), 1067–1071.

Henze, K., Badr, A., Wettern, M., Cerff, R., Martin, W., 1995. A nuclear gene of eubacterial origin in *Euglena gracilis* reflects cryptic endosymbioses during protist evolution. Proc. Natl. Acad. Sci. U. S. A. 92, 9122–9126.

Kucharski, R., Maleszka, R., 2002. Evaluation of differential gene expression during behavioral development in the honeybee using microarrays and Northern blots. Genome Biol. 3 (2) (research 0007, 1-0007, 9).

Leander, B., 2004. Did trypanosomatid parasites have photosynthetic ancestors? Trends Microbiol. 12 (6), 251–258.

Linton, E.W., Nudelman, M.A., Conforti, V., Triemer, R.E., 2000. A molecular analysis of the euglenophytes using SSU rDNA. J. Phycol. 36, 740–746.

Manusadzianas, L., Maksimov, G., Darginaviciene, J., Jurkoniene, S., Sadauskas, K., Vitkus, R., 2002. Response of the chlorophyte *Nitellopsis obtusa* to heavy metals at the cellular, cell membrane and enzyme levels. Environ. Toxicol. 17 (3), 275–283.

Margulis, L., 1993. Symbiosis in cell evolution. Microbial Communities in the Archean and Proterozoic Eons. W. Freeman and Company, New York, p. 452.

Marin, B., Palm, A., Klingberg, M., Melkonian, M., 2003. Phylogeny and taxonomic revision of plastid-containing euglenophytes based on SSU rDNA sequence comparisons and synapomorphic signatures in the SSU rRNA secondary structure. Protist 154, 99–145.

Maslov, D.A., Yasuhira, S., Simpson, L., 1999. Phylogenetic affinities of *Diplonema* within Euglenozoa as inferred from SSU rRNA gene and partial COI protein sequences. Protist 150 (1), 33–42.

McCord, J.M., Fridovich, I., 1969. Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J. Biol. Chem. 244 (22), 6049–6055.

Müllner, A.N, Angler, D.G., Samuel, R., Linton, E.W., Triemer, R.E., 2001. Phylogenetic analysis of phagotrophic, phototrophic, and osmotrophic euglenoids by using the nuclear 18S rDNA sequence. Int. J. Syst. Evol. Microbiol. 51, 783–791.

Nozaki, H., Ohta, N., Matsuzaki, M., Misumi, O., Kuroiwa, T., 2003. Phylogeny of plastids based on cladistic analysis of gene loss inferred from complete plastid genome sequences. J. Mol. Evol. 57, 377–382.

Nudelman, M.A., Rossi, M.S., Conforti, V., Triemer, R.E., 2003. Phylogeny of *Euglenophyceae* based on small subunit rDNA sequences: taxonomic implications. J. Phycol. 39, 226–235.

Ohlrogge, J.B., 1997. Regulation of fatty acid synthesis. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48, 109–136.

Parkinson, J., Guiliano, D.B., Blaxter, M., 2002. Making sense of EST sequences by CLOBBing them. BMC Bioinformatics 3, 31.

Ramdas, L., et al., 2001. Sources of nonlinearity in cDNA microarrays expression measurements. Genome Biol. 2 (research 0047. 1-0047.7).

Rocchetta, I., Ruiz, L.B., Magaz, G., Conforti, V.T., 2003. Effects of hexavalent Chromium in two strains of *Euglena gracilis*. Bull. Environ. Contam. Toxicol. 70 (5), 1045–1051.

Romano, P., Horton, P., Gray, J.E., 2004. The *Arabidopsis* cyclophilin gene family. Plant Physiol. 134 (4), 1268–1282.

Rosenberg, A., 1976. Light-independent stoichiometry of galactosyl diglyceride and chlorophyll accretion during light-induced chloroplast membrane synthesis in *Euglena*. Biochem. Biophys. Res. Commun. 73 (4), 972–977.

Ruiz, L.B., Rocchetta, I., dos Santos Ferreira, V., Conforti, V., 2004. Isolation, culture and characterization of a new strain of *Euglena gracilis*. Phycol. Res. 52, 168–179.

Russell, A.G., Watanabe, Y., Charette, J.M., Gray, M., 2005. Unusual features of fibrillarin cDNA and gene structure in *Euglena gracilis*: evolutionary conservation of core proteins and structural predictions for methylation-guide box C/D snoRNPs throughout the domain Eucarya. Nucleic Acids Res. 33 (9), 2781–2791.

Schawalder, S.B., Kabani, M., Howald, I., Choudhury, U., Werner, M., Shore, D., 2004. Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh 1. Nature 432 (7020), 1058–1061.

Schiff, J.A., Schwartzbach, S.D., 1982. Photocontrol of chloroplast development in *Euglena*. The Biology of *Euglena*. . Physiology, vol. III. Academic Press, pp. 313–352.

Schwartzbach, S.D., Schiff, J., 1974. Chloroplast and cytoplasmic ribosomes of *Euglena*: selective binding of dihydrostreptomycin to chloroplast ribosomes. J. Bacteriol. 120 n1, 334–341.

Triemer, R.E., et al., 2006. Phylogeny of the Euglenales based upon combined SSU and LSU rDNA sequence comparisons and description of *Discoplastis* Gen. Nov. (Euglenophyta). J. Phycol. 42 (3), 731–740.

Yasuhira, S., Simpson, L., 1997. Phylogenetic affinity of mitochondria of *Euglena gracilis* and kinetoplastids using cytochrome oxidase I and hsp60. J. Mol. Evol. 44 (3), 341–347.