

# Evolution and diversity of glutaredoxins in photosynthetic organisms

Jérémy Couturier · Jean-Pierre Jacquot ·  
Nicolas Rouhier

Received: 9 April 2009 / Revised: 6 May 2009 / Accepted: 19 May 2009 / Published online: 9 June 2009  
© Birkhäuser Verlag, Basel/Switzerland 2009

**Abstract** The genome sequencing of prokaryotic and eukaryotic photosynthetic organisms enables a comparative genomic study of the glutaredoxin (Grx) family. The analysis of 58 genomes, using a specific motif composed of the active site sequence and of amino acids involved in glutathione binding, led to an updated classification of Grxs into six classes. Only two classes (I and II) are common to all photosynthetic organisms. Eukaryotes and cyanobacteria have two specific Grx classes (classes III and IV and classes V and VI, respectively). The classes IV, V and VI have not yet been identified and contain multimodular Grx fusions. In addition, putative Grx partners were identified from the presence of fusion proteins, the conservation of gene order in bacterial operons, and the gene co-occurrence. The genes encoding class II Grxs and BoLA/YrbA proteins are frequently adjacent, in the same transcriptional orientation in prokaryote genomes and present in the same organisms.

**Keywords** Evolution · Genomic · Glutaredoxin · Photosynthetic organisms

## Introduction

Glutaredoxins (Grxs) are glutathione (GSH)- or thioredoxin reductase (TR)-dependent oxidoreductases which are conserved in most eukaryotes and prokaryotes, except in some bacterial or archaeal phyla [1, 2]. When Grxs are recycled by GSH, the GSSG formed is in turn reduced by NADPH and glutathione reductase (GR) forming the GSH/Grx reducing system. Alternatively, many Grxs from various organisms are recycled by ferredoxin- or NADPH-dependent thioredoxin reductases forming a TR/Grx reducing system [1, 3–5]. Grxs share a conserved 3D structure with members of the thioredoxin (TRx) superfamily and function in the reduction of disulfide bonds, especially with glutathionylated proteins.

Grxs were initially categorized, based on the active site sequence, into two groups, a dithiol (CPY/FC motif) and a monothiol (CGFS motif) subgroup [6]. The dithiol Grxs were discovered a long time ago in *Escherichia coli* as an alternative donor to Trxs for ribonucleotide reductase. Two of the Grxs from *E. coli* (Grx1 and 3) are small proteins (around 85 amino acids) with a very classical YCPYC active site which is likely to be the ancestral sequence from which most other Grxs from class I diverged. In contrast, the CGFS Grxs were discovered very recently (late 1990s) and are only present in eukaryotes, in most proteobacteria except in the campylobacterale order, and in cyanobacteria but not in other bacterial phyla and archaea except in the halobacteriale order. However, the recent sequencing of many genomes has led to the identification of a broader diversity among these sequences. In particular, in land plants, Grx isoforms have been classified into three distinct classes based essentially on the overall primary sequence and on their active site structures [2, 7, 8]. Class I includes proteins with CxxC/S active sites other than CGFS. It is the

---

**Electronic supplementary material** The online version of this article (doi:10.1007/s00018-009-0054-y) contains supplementary material, which is available to authorized users.

---

J. Couturier · J.-P. Jacquot · N. Rouhier (✉)  
Interactions Arbres Microorganismes, IFR 110 Génomique  
Ecophysiologie et Ecologie Fonctionnelles,  
Unité Mixte de Recherches 1136 INRA-Nancy Université,  
54506 Vandœuvre-lès-Nancy Cedex, France  
e-mail: nrouhier@scbiol.uhp-nancy.fr

most widespread class, and organisms possessing Grxs from this class generally have one to six isoforms. Class II contains Grxs with a CGFS motif (with a few exceptions), and organisms which possess Grxs of this class have generally one to six isoforms. Class III, specific to plants, corresponds to Grxs with a peculiar CCxx active site, very often CCxC or CCxS. The large number of Grxs in this class explains why the Grx family is expanded in terrestrial plants [8]. Although these proteins have very diversified active sites, divergent from the other two classes, they conserved the amino acids involved in glutathione recognition and shared by all Grxs.

Due to the existence of mono- and di-thiol isoforms, Grxs can use several catalytic mechanisms. For the reduction of intramolecular disulfide bonds on target proteins, the mechanism defined as a dithiol type mechanism is similar to that of the thioredoxins. The first Grx active site cysteine attacks the target disulfide and a mixed disulfide is formed between the two proteins. Then, the second active site cysteine is required for solving this intermediate disulfide bridge. Nevertheless, in most cases, Grxs rather reduce specifically protein–glutathione adducts via two distinct mechanisms. In both cases, the N-terminus active site cysteine is employed for reducing mixed disulfides between GSH and the target proteins, leading to a glutathionylated Grx. In the monothiol mechanism, one molecule of glutathione is used to directly regenerate the glutathionylated Grxs. In the dithiol mechanism, a resolving cysteine is needed. It can be either the second active site cysteine, as shown for human Grx2, or a conserved extra-active site C-terminus cysteine, as suggested for yeast Grx5 and CrGrx3 based on the biochemical evidence that a disulfide bond is formed between the two cysteines [1, 3, 5, 9]. In the latter two cases, this leads to the formation of an intramolecular disulfide bridge which is reduced either by GSH or TRs. Hence, the situation is complicated by the fact that Grxs with a dithiol active site can employ both monothiol and dithiol mechanisms, and that Grxs with a monothiol active site do the same if an extra active site cysteine can serve as a resolving cysteine.

This paper will not review the physiological functions of Grxs in photosynthetic organisms as this has been done recently [10–12], but it will present an update of the Grx classification. Previous comparative genomics analyses performed in three higher plants (*Arabidopsis thaliana*, *Populus trichocarpa* and *Oryza sativa*), in the alga *Chlamydomonas reinhardtii* and in the cyanobacteria *Synechocystis* sp. PCC 6308 highlighted the presence of expanded families in higher plants (from 27 to 35 genes compared to 6 and 3 genes in green algae and cyanobacteria, respectively), whereas non-photosynthetic organisms contain only a limited number of these genes

and proteins [2, 7, 8]. The increasing number of sequenced genomes from photosynthetic organisms reveals that a clear and global classification of the Grx family is still not fully accomplished. This is also true in other kingdoms, which contain a lot of non-annotated and uncharacterized Grx and Grx-like proteins [13]. Here, we have used the most recent genomic data to decipher the diversity and evolution of Grxs in photosynthetic organisms. The present classification has been refined using genomic sequences from 58 organisms, 7 vascular plants (4 dicots, 2 monocots, one lycophyte), 1 non-vascular plant (a bryophyte), 12 algae (8 unicellular or multicellular green algae, 2 diatoms, 1 red alga, 1 haptophyte) and 38 cyanobacteria. Compared to the earlier classification, there is an additional Grx class (class IV) in eukaryotes consisting of proteins with three domains, an N-terminal Grx module followed by two domains of unknown function. In addition, there are two Grx classes, restricted to cyanobacteria, composed of elongated Grxs containing either an N-terminal domain of unknown function or a C-terminal transmembrane portion (classes V and VI, respectively). Members of class V are present in some other bacteria, whereas class VI appears to be restricted to cyanobacteria.

## Materials and methods

Bioinformatic genome analysis: sequence annotation, phylogenetic analyses

The Grx sequences retrieved by text and Blast searches from the *P. trichocarpa* whole genome database (version 1.1) at the U.S. Department of Energy Joint Genome Institute (JGI) ([http://genome.jgi-psf.org/Poptr1\\_1/Poptr1\\_1.home.html](http://genome.jgi-psf.org/Poptr1_1/Poptr1_1.home.html)) have been previously corrected [8]. The curated poplar amino acid sequences were used to search against 57 other genomes from photosynthetic organisms using BLASTP or TBLASTN. The genomes are available at the following websites, for *A. thaliana* (<http://www.arabidopsis.org/>), *O. sativa* (<http://rice.plantbiology.msu.edu/>), *Vitis vinifera* (<http://www.genoscope.cns.fr/spip/Vitis-vinifera-whole-genome.html>), *Sorghum bicolor* (<http://genome.jgi-psf.org/Sorbi1/Sorbi1.home.html>), *Medicago truncatula* (version 2.0) (<http://mips.gsf.de/proj/plant/jsf/medi/index.jsp>), *C. reinhardtii* (version 3.0) (<http://genome.jgi-psf.org/Chlre3/Chlre3.home.html>), *Ostreococcus lucimarinus* (version 2.0) ([http://genome.jgi-psf.org/Ost9901\\_3/Ost9901\\_3.home.html](http://genome.jgi-psf.org/Ost9901_3/Ost9901_3.home.html)), *Ostreococcus RCC809* (version 1.0) ([http://genome.jgi-psf.org/OstRCC809\\_1/OstRCC809\\_1.home.html](http://genome.jgi-psf.org/OstRCC809_1/OstRCC809_1.home.html)), *Ostreococcus tauri* (version 2.0) (<http://genome.jgi-psf.org/Osta4/Osta4.home.html>), *Physcomitrella patens* subsp. *patens* (version 1.1) (<http://genome.jgi-psf.org/>

[Phypa1\\_1/Phypa1\\_1.home.html](#)), *Selaginella moellendorffii* (version 1.0) ([http://genome.jgi-psf.org/Selmo1/Selmo1.home.html](#)), *Chlorella vulgaris* C-169 (version 1.0) ([http://genome.jgi-psf.org/Chlvu1/Chlvu1.home.html](#)), *Volvox carteri* (version 1.0) ([http://genome.jgi-psf.org/Volca1/Volca1.home.html](#)), *Thalassiosira pseudonana* ([http://genome.jgi-psf.org/Thaps3/Thaps3.home.html](#)), *Phaeodactylum tricornutum* ([http://genome.jgi-psf.org/Phatr2/Phatr2.home.html](#)), *Cyanidioschyzon merolae* ([http://merolae.biol.s.u-tokyo.ac.jp/blast/blast.html](#)), *Emiliania huxleyi* CCMP1516 ([http://genome.jgi-psf.org/Emihu1/Emihu1.home.html](#)), *Micromonas pusilla* CCMP1545 ([http://genome.jgi-psf.org/MicpuC2/MicpuC2.home.html](#)), *Micromonas* strain RCC299 ([http://genome.jgi-psf.org/MicpuN2/MicpuN2.home.html](#)), and 38 cyanobacteria ([http://bacteria.kazusa.or.jp/cyanobase/](#)) (see Table 1 for the complete list). Whenever it was possible, all the incomplete sequences have been correctly annotated based on available ESTs and manual inspection of the genomic sequences. In addition to EST sequences available in Genbank, some ESTs have been retrieved from the DFCI database ([http://compbio.dfci.harvard.edu/tgi/plant.html](#)). All protein sequences and corresponding accession numbers used in this article can be found in the databases mentioned above and as electronic supplementary material (ESM). The amino acid sequence alignments were done using CLUSTALW and imported into the Molecular Evolutionary Genetics Analysis (MEGA) package version 4.1 [14]. Phylogenetic analyses were conducted using the neighbor-joining (NJ) method implemented in MEGA, with the pairwise deletion option for handling alignment gaps, and with the Poisson correction model for distance computation. Bootstrap tests were conducted using 1,000 replicates. Branch lengths are proportional to phylogenetic distances.

Analysis of gene fusion, gene clustering,  
and gene co-occurrence

The gene order comparison and clustering of *grx* genes in bacterial genomes has been analyzed using the “protein clusters” tool available at the NCBI webpage and the Microbial Genome Database (MGDB, [http://mbgd.genome.ad.jp/](#)). The STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) webtool ([http://string.embl.de/](#)) enabled the detection of the existence of fusion proteins, of gene co-occurrence and more generally the prediction of putative protein–protein interactions [15]. Class I, V and VI Grxs are included in the COG0695 (cluster of orthologous groups), whereas class II Grxs are present in the COG0278. The *bolA* members identified in this genome analysis belong to COG0271.

## Results and discussion

The specific traits of the selected photosynthetic organisms

For this in silico comparative genomic analysis, we have selected several completely sequenced photosynthetic organisms, which possess different ways of life and thus constitute models for studying the evolution of photosynthesis, vascular tissue development, flowering and wood formation. Cyanobacteria represent prokaryotic organisms which obtain their energy through photosynthesis. All other organisms are unicellular or multicellular eukaryotes. Regarding the algal kingdom, we have selected a rhodophyte (the red alga *Cyanidioschyzon merolae*), 8 chlorophytes belonging to different classes and genus, and 3 Chromoalveolata, i.e., 2 heterokontophyta (the diatoms *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*) and a haptophyta, *Emiliania huxleyi* (ESM, Fig. 1). *Ostreococcus* and *Micromonas* species belong to the Prasinophyceae, an early-diverging class within the green lineage, while *C. reinhardtii*, *V. carteri* and *C. vulgaris* C169 belong to the Chlorophyceae. Regarding *Ostreococcus*, the genomes of three different species, which differed in their adaptation to variable light intensity, are now available. The comparison between *V. carteri*, a multicellular organism, and the single-celled *C. reinhardtii* and *C. vulgaris* should provide insights into the evolution of multicellularity in the green algal lineage and beyond. The two diatoms (*T. pseudonana* and *P. tricornutum*) and the coccolithophore *E. huxleyi* represent photosynthetic microorganisms found throughout marine and freshwater ecosystems. The comparison with green algae could thus bring information on the differences between marine and freshwater environments. Then, the passage from water to terrestrial life can be illustrated by investigating land plants. *Physcomitrella patens* ssp. *patens* (later referred to as *P. patens*, uniquely) is a moss species, which stands at the base of terrestrial plants, having diverged before the acquisition of well-developed vascular tissues. The lycophyte, *Selaginella moellendorffii*, constitutes an ancient group of plants, which sits phylogenetically between bryophytes and higher plants. Hence, lycophytes are at a key position to understand the appearance of vascular tissues. Together with bryophytes, they can provide information on how major innovations evolved in order for plants to survive and thrive on land. Other organisms selected are flowering plants belonging to the dicot and monocot families. The two monocots possess different photosynthetic pathways for carbon assimilation. *Sorghum bicolor* is a representative of tropical grasses, which use ‘C<sub>4</sub>’

**Table 1** Gene content and distribution of cyanobacterial glutaredoxins

|   | Class I |         | Class II | Class V | Class VI | Other | Total |
|---|---------|---------|----------|---------|----------|-------|-------|
|   | CxxC/S  | Prx/Grx | CGFS     | CPWG    | CPWS/C   |       |       |
| <i>Thermosynechococcus elongatus</i> BP-1 <sup>a</sup>            | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Gloeobacter violaceus</i> PCC 7421 <sup>b</sup>                | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Synechococcus elongatus</i> PCC 7942 <sup>a</sup>              | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Synechococcus elongatus</i> PCC 6301 <sup>a</sup>              | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Synechococcus</i> sp. JA-2-3B <sup>a</sup> (2-13) <sup>a</sup> | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Synechococcus</i> sp. JA-3-3Ab <sup>a</sup>                    | 1       | 0       | 1        | 0       | 0        | 0     | 2     |
| <i>Synechococcus</i> sp. RCC307 <sup>a</sup>                      | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Synechococcus</i> sp. WH8102 <sup>a</sup>                      | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Synechococcus</i> sp. CC9311 <sup>a</sup>                      | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Synechococcus</i> sp. CC9902 <sup>a</sup>                      | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Synechococcus</i> sp. SYNCC9605 <sup>a</sup>                   | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. NATL2A <sup>c</sup>           | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> MED4 <sup>c</sup>                  | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> MIT9313 <sup>c</sup>               | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9312 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. AS9601 <sup>c</sup>           | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9515 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9303 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. NATL1A <sup>c</sup>           | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9301 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9215 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> str. MIT 9211 <sup>c</sup>         | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Prochlorococcus marinus</i> SS120 <sup>c</sup>                 | 1       | 0       | 1        | 0       | 1        | 0     | 3     |
| <i>Cyanothece</i> sp. PCC 7425 <sup>a</sup>                       | 1       | 0       | 1        | 1       | 0        | 0     | 3     |
| <i>Synechocystis</i> sp. PCC 6803 <sup>a</sup>                    | 2       | 0       | 1        | 0       | 0        | 0     | 3     |
| <i>Synechococcus</i> sp. PCC 7002 <sup>a</sup>                    | 2       | 0       | 1        | 0       | 0        | 0     | 3     |
| <i>Cyanothece</i> sp. ATCC 51142 <sup>a</sup>                     | 2       | 0       | 1        | 0       | 0        | 0     | 3     |
| <i>Anabaena variabilis</i> ATCC 29413 <sup>d</sup>                | 2       | 1       | 1        | 0       | 0        | 0     | 4     |
| <i>Anabaena</i> sp. PCC 7120 <sup>d</sup>                         | 2       | 1       | 1        | 0       | 0        | 0     | 4     |
| <i>Nostoc punctiforme</i> ATCC 29133 <sup>d</sup>                 | 2       | 1       | 1        | 0       | 0        | 0     | 4     |
| <i>Microcystis aeruginosa</i> NIES-843 <sup>a</sup>               | 2       | 1       | 1        | 0       | 0        | 0     | 4     |
| <i>Synechococcus</i> sp. WH 7803 <sup>a</sup>                     | 1       | 0       | 1        | 1       | 1        | 0     | 4     |
| <i>Synechococcus</i> sp. WH 7805 <sup>a</sup>                     | 1       | 0       | 1        | 1       | 1        | 0     | 4     |
| <i>Synechococcus</i> sp. RS9917 <sup>a</sup>                      | 1       | 0       | 1        | 2       | 1        | 0     | 5     |
| <i>Cyanobium</i> sp. PCC 7001 <sup>a</sup>                        | 1       | 0       | 1        | 2       | 1        | 0     | 5     |
| <i>Cyanothece</i> sp. PCC 8801 <sup>a</sup>                       | 2       | 1       | 1        | 1       | 0        | 0     | 5     |
| <i>Cyanothece</i> sp. PCC 8802 <sup>a</sup>                       | 2       | 1       | 1        | 1       | 0        | 0     | 5     |
| <i>Acaryochloris marina</i> MBIC11017 <sup>e</sup>                | 4       | 0       | 1        | 1       | 0        | 1     | 7     |
| Total   | 50      | 6       | 38       | 10      | 21       | 1     | 126   |

The Grx sequences have been identified in 38 different cyanobacterial genomes. They essentially belong to four different classes. Strikingly, the total number of Grx isoforms varies from 2 to 7, depending on the species considered, a value comparable to non-photosynthetic eukaryotes. *A. marina* MBIC11017 possesses a unique Grx, not found in other species, except some  $\gamma$ -proteobacteria. While class I and II Grxs are universally distributed, the Grxs from classes V or VI are restricted to some species

<sup>a</sup> Cyanobacteria of the order Chroococcales

<sup>b</sup> Cyanobacteria of the order Gloeobacterales

<sup>c</sup> Cyanobacteria of the order Prochlorococcales

<sup>d</sup> Cyanobacteria of the order Nostocales

<sup>e</sup> Unclassified cyanobacteria

photosynthesis and has thus developed particular biochemical and morphological specializations to improve carbon assimilation at high temperatures. In contrast, rice is a representative of temperate grasses, using 'C<sub>3</sub>' photosynthesis. *M. truncatula* is a model dicot organism for legume biology. It forms arbuscular mycorrhiza with fungi and symbioses with nitrogen-fixing rhizobia. The model plant *A. thaliana* does not form symbiosis naturally, making *M. truncatula* an important tool for studying these processes. Perennial liana, such as *V. vinifera*, have many developmental aspects, which make them different from other dicots, including a unique shoot architecture with petiolated leaves opposite to either inflorescences or tendrils. *Populus* is considered as a model of woody plant, especially for studying tree-specific traits such as wood formation or perennial growth.

#### Identification and classification of mono- and multi-modular Grxs

Using previously annotated Grxs from *P. trichocarpa*, *A. thaliana* and *O. sativa*, we performed, both by text and blastp or tblastn searches, a complete analysis of the Grx family in a large number of photosynthetic organisms. During this process, many truncated or mis-annotated Grx sequences have been corrected and annotated for the reliability of this analysis. A total of 428 sequences, most of them complete, are now available in the ESM.

For the classification of this improved set of Grxs, we have considered several features. First, except for a few proteins (see below), they generally possess a CxxC/S active site motif situated at the N-terminus of an alpha helix. The previous classification of plant Grxs was based on the presence of either a cysteine or a serine at the fourth position of the active-site. Among the 31 *Arabidopsis* Grxs, 14 Grxs possess a cysteine and have been named Grx C1 to C14, and 17 Grxs possess a serine and have been called Grx S1 to S17 [2]. Second, we have used amino acids determined, from structural studies, as involved in glutathione binding. In particular, there is a TVP sequence (or any analogous motif) located ca. 35–40 amino acids after the active site, and containing the proline conserved in the oxidoreductases of the Trx superfamily, such as thioredoxins, protein disulfide isomerases, and bacterial Dsb proteins, and usually found in the *cis* conformation. Around ten amino acids after the TVP motif, there are two conserved glycines forming a structural kink at the proximity of the active site. As an example, the recently structurally characterized GrxS12 exhibits the following motif <sup>29</sup>CSYS<sub>32</sub>, <sup>73</sup>TVP<sub>75</sub> and <sup>85</sup>GG<sub>86</sub> [16]. Hence, excluding targeting sequences, we defined a Grx domain as a protein sequence composed of 90–120 amino acids and containing

the characteristic sequences detailed above. We have thus considered any sequence possessing at least one Grx domain, including large proteins which exhibited additional protein domains.

The Grx sequences retrieved from 38 cyanobacterial and 20 eukaryote genomes were used for phylogenetic analysis (Tables 1, 2). Six Grx classes in photosynthetic organisms have been identified but, as classes I and II are common to all organisms, cyanobacterial and eukaryote sequences can both be classified into four classes. Additionally, we have detected two particular Grx sequences from *Acaryochloris marina* MBIC11017 (a Grx of 205 amino acids, elongated in the N-terminal part and only found in some  $\gamma$ -proteobacteria of the vibrio genus) and from *C. reinhardtii* (a fusion of 940 amino acids with a C-terminal domain homologous to a dicarboxylate transporter). They are not clustering into one of the six classes and they have no orthologs in other photosynthetic organisms. Regarding the number of members and their distribution in the different classes, some important differences exist between all photosynthetic organisms. First, approximately 30 different Grx isoforms can be identified in higher plants (Table 2). Other land plants (mosses and lycophytes) apparently possess around 15 isoforms, but with a different distribution in classes II and III compared to higher plants. The most striking example is the presence of only 2 Grxs from class III in these lower plants while there are between 13 and 24 Grxs in higher plants. It illustrates the intermediate evolutionary position of these organisms and suggests that the expansion of this class correlates in particular with the appearance of flowering capacity. In algae, the lower number of Grx isoforms, ranging from 4 to 10, is essentially explained by the absence of class III Grxs. In cyanobacteria, if we do not consider the specific case of *Acaryochloris marina* MBIC11017, the number of Grx members comprises between 2 and 5 (Table 1). All cyanobacteria possess at least one member of class I and II, and the differences arise from the presence of hybrid proteins composed of an N-terminal peroxiredoxin (Prx) module fused to a C-terminal Grx module and named PrxGrx fusions (which cluster into class I), and of two additional classes present only in some species.

Figure 1 shows a phylogenetic tree with most Grx sequences from the organisms analyzed. Although the sequences from cyanobacteria, algae and moss are more distantly related from angiosperms and very often constitute separate branches, they usually fall in the right class. In this new phylogenetic classification, the distribution of the Grxs previously identified from higher plants is still consistent with the former classification (classes I to III). Nevertheless, for class I, the cyanobacterial sequences cluster separately from eukaryote Grxs as they are quite divergent in terms of primary sequence (Figs. 1, 2).



**Table 2** Gene content and distribution of glutaredoxins in eukaryote photosynthetic organisms

|             |                                    | Class I |    |    |    |    | Class II   |       |     |     |     | Class III | Class IV | Other | Total |
|-------------|------------------------------------|---------|----|----|----|----|------------|-------|-----|-----|-----|-----------|----------|-------|-------|
|             |                                    | Total   | C1 | C2 | C3 | C4 | C5/<br>S12 | Total | S14 | S15 | S16 | S17       |          |       |       |
| Haptophyte  | <i>Emiliania huxleyi</i>           | 6       | –  | –  | –  | –  | –          | 3     | 1   | 1   | 0   | 1         | 0        | 0     | 9     |
| Diatoms     | <i>Phaeodactylum tricornutum</i>   | 4       | –  | –  | –  | –  | –          | 3     | 1   | 1   | 0   | 1         | 0        | 0     | 6     |
|             | <i>Thalassiosira pseudonana</i>    | 2       | –  | –  | –  | –  | –          | 3     | 1   | 1   | 0   | 1         | 0        | 1     | 5     |
| Red algae   | <i>Cyanidioschyzon merolae</i>     | 2       | –  | –  | –  | –  | –          | 2     | 1   | 1   | 0   | 0         | 0        | 0     | 4     |
| Green algae | <i>Micromonas pusilla CCMP1545</i> | 3       | –  | –  | –  | –  | –          | 5     | 2   | 1   | 1   | 1         | 0        | 2     | 10    |
|             | <i>Micromonas RCC299</i>           | 4       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 2     | 10    |
|             | <i>Ostreococcus tauri</i>          | 0       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 5     |
|             | <i>Ostreococcus lucimarinus</i>    | 0       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 5     |
|             | <i>Ostreococcus RCC809</i>         | 0       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 5     |
|             | <i>Chlorella vulgaris</i>          | 2       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 7     |
|             | <i>Volvox carteri</i>              | 2       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 7     |
|             | <i>Chlamydomonas reinhardtii</i>   | 2       | –  | –  | –  | –  | –          | 4     | 1   | 1   | 1   | 1         | 0        | 1     | 8     |
| Moss        | <i>Physcomitrella patens</i>       | 5       | 0  | 3  | 1  | 0  | 1          | 8     | 2   | 2   | 1   | 3         | 2        | 0     | 15    |
| Lycophytes  | <i>Selaginella moellendorffii</i>  | 4       | 0  | 2  | 0  | 1  | 1          | 9     | 3   | 3   | 1   | 2         | 2        | 1     | 16    |
| Monocots    | <i>Sorghum bicolor</i>             | 5       | 0  | 2  | 1  | 1  | 1          | 6     | 1   | 2   | 2   | 1         | 19       | 2     | 32    |
|             | <i>Oryza sativa</i>                | 5       | 0  | 2  | 1  | 1  | 1          | 5     | 1   | 2   | 1   | 1         | 17       | 2     | 29    |
| Dicots      | <i>Vitis vinifera</i>              | 5       | 1  | 1  | 1  | 1  | 1          | 5     | 2   | 1   | 1   | 1         | 13       | 2     | 25    |
|             | <i>Medicago truncatula</i>         | 6       | 2  | 1  | 1  | 1  | 1          | 5     | 2   | 1   | 1   | 1         | 18       | 2     | 31    |
|             | <i>Arabidopsis thaliana</i>        | 6       | 1  | 1  | 1  | 1  | 2          | 4     | 1   | 1   | 1   | 1         | 21       | 2     | 33    |
|             | <i>Populus trichocarpa</i>         | 6       | 2  | 1  | 1  | 1  | 1          | 5     | 1   | 1   | 1   | 2         | 24       | 3     | 38    |
| Total       |                                    | 69      |    |    |    |    |            | 91    |     |     |     |           | 116      | 25    | 302   |

The Grx sequences have been identified in 20 different eukaryote genomes (algae, bryophyte, lycophyte, and angiosperms). They essentially belong to four different classes, *C. reinhardtii* having a specific fusion protein (a Grx domain followed by a domain with a homology to a dicarboxylate transporter) not found elsewhere. As for cyanobacteria, all organisms, except *Ostreococcus* species, contain Grxs from classes I and II. Grxs of class III are only found in land plants and those of class IV in most organisms

Moreover, there are PrxGrx fusion proteins in cyanobacteria containing a Grx domain belonging to class I. On the contrary, cyanobacterial CGFS Grxs are clustering with eukaryotic sequences (Fig. 1).

The remaining sequences define three new classes containing multi-modular Grxs, not identified before as Grxs. Class IV, specific to eukaryotes, contains proteins with very diverse CxxC or CxxS (in a few cases) active sites. Classes V and VI (containing proteins with CPWG and CPWC/S active sites, respectively) are specific to cyanobacteria. A detailed analysis of each class will be presented below.

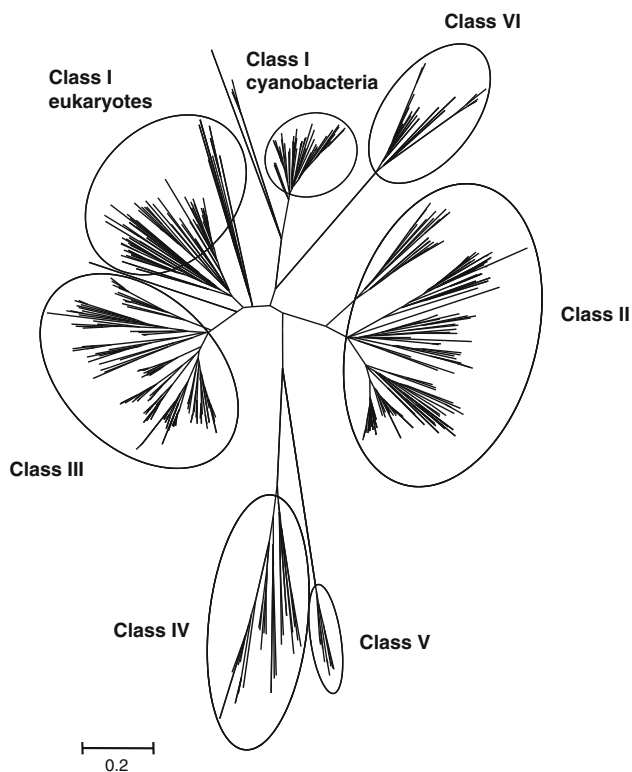
#### The ubiquitous class I glutaredoxins

##### Overall view

The phylogenetic analysis revealed two major subgroups, containing either cyanobacterial or eukaryote sequences, with a few sequences isolated coming from diatoms, haptophyte or *Micromonas*, most likely arising from the lower representativity of these organisms (Fig. 2). From

the 119 cyanobacterial and eukaryote Grx sequences used in this analysis, a motif specific to class I Grxs has been defined as **C**[P/G/S/D][F/Y]<sub>H/P/W/R</sub>[C/S/T]X<sub>35</sub>[T/S/R][V/I/L/M]**PX<sub>8</sub>GG**, with four invariant residues (shown in bold). The lowercase letter represents poorly used amino acids and X the number of amino acids separating the different parts of the motif. This class contains orthologs of the classical dithiol Grxs already described in other organisms (*E. coli* Grx1 to 3, yeast and human Grx1 and 2).

All cyanobacteria analyzed contain at least one class I Grx, but surprisingly there are 4 members in *A. marina* (Table 1). All cyanobacterial class I Grxs display a conserved CPYC or CPFC dithiol active site, but they are clustering into at least five different subgroups. In species (belonging to Chroococcales or Nostocales) that contain two or more genes, it may arise from an ancient duplication event as the two genes are in general quite divergent, from 57.9 to 66.7% identity. As a matter of fact, they are forming a single branch (subgroup 1). Subgroup 2 is constituted by sequences from the *Prochlorococcus* genus. Subgroup 3 is composed of the so-called PrxGrx (Figs. 1, 2). Bacterial orthologs have been identified and



**Fig. 1** Unrooted, neighbor-joining (NJ)-based tree of the Grx family in photosynthetic organisms. The analysis was performed using MEGA 4 with the setup described in “Material and methods”. Branch lengths are proportional to phylogenetic distances. For clarity, the protein names have been removed but they are available in Figs. 3, 4, 5, 6 and 7. The full names and the corresponding accession numbers of all Grx sequences are available in ESM file 1. Around 99% of the members are clustering into one of the six Grx classes defined

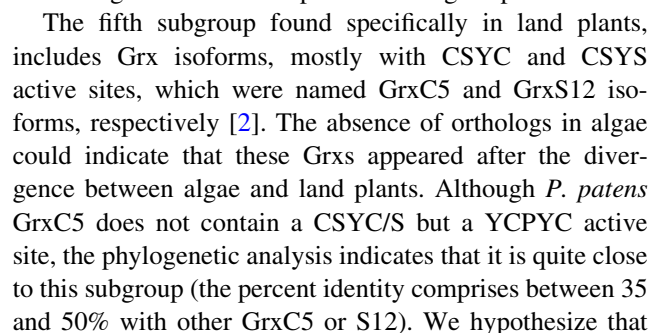
characterized in various pathogenic bacteria such as *Haemophilus influenzae* and *Neisseria meningitidis* [17, 18]. In *Anabaena* sp. PCC7120, this protein also possesses, as expected, a peroxidase activity toward hydrogen peroxide using reduced glutathione as an electron donor and acts as an important player in hydrogen peroxide detoxification in late-phase growth [19]. These authors concluded that this protein would be only present in Nostocales, but this genomic analysis indicates that there is one member in the chroococcale *Microcystis aeruginosa* NIES-843. No PrxGrx orthologs have been retrieved from photosynthetic eukaryotes genomes suggesting that, during evolution, the gene encoding these fusion proteins has either been lost or has been separated into two distinct Prx and Grx genes. With the exception of *A. marina*, whose classification is still as matter of debate, the two last subgroups (4 and 5) contain Grxs from both Chroococcales and Nostocales.

The Grx sequences from higher plants, either dicots or monocots, belong to the five previously identified subgroups called GrxC1, C2, C3, C4 and C5/S12, with GrxC1 being more related to GrxC2 and GrxC3 being more related

to GrxC4 (Fig. 2) [8]. The sequences from the moss *P. patens* and from the lycophyte *S. moellendorffii* do not always perfectly integrate into these groups, although some are clearly related to a given subgroup. For example, from their active site and their position on the phylogenetic tree, Grxs from *P. patens* and *S. moellendorffii* (PpGrxC5 and SmGrxC5 or PpGrxC2.1 to C2.3 and SmGrxC2.1 and C2.2) are clustering, respectively, with the higher plant GrxC5/S12 and GrxC2 subgroups. On the contrary, the classification is clearly not adapted to algae. Most of the time, they are grouping in separate branches, which do not fit with their phylum, or are dispersed. Hence, depending on the organisms and based on their identity, they have been named arbitrarily from Grx1.1 to Grx1.4 and from Grx2.1 to Grx2.3, following the nomenclature used in *C. reinhardtii* where Grx1 and 2 are dithiol Grxs and Grx3 to 6 represent CGFS Grxs. Overall, the class I Grx has been expanded in land plants during evolution since the number of Grx ranges from 4 to 6 in land plants and from 0 to 4 in algae, except *E. huxleyi* which contains 6 isoforms most likely arising from specific duplication events (Table 2). The absence of class I Grxs in the three sequenced *Ostreococcus* species is striking, considering that there are 3 or 4 class I Grxs in *Micromonas*, another genus of the Mamiellaceae family. The most plausible explanation is that they have been lost specifically in the *Ostreococcus* genus during the evolution process.

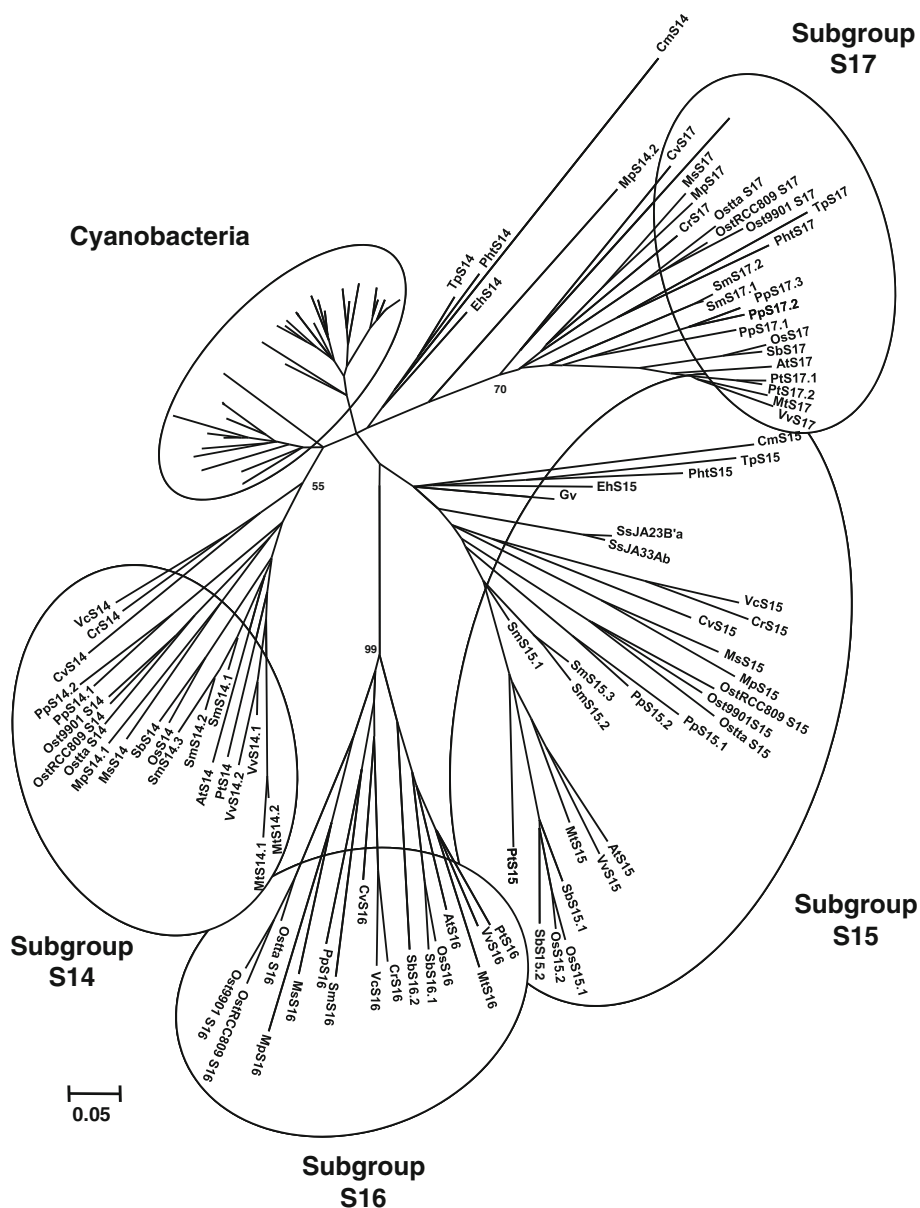
#### A detailed analysis of the higher plant subgroups

Owing to the higher number of representatives in land plants, each subgroup exhibits some specific features. The GrxC1 subgroup contains isoforms with a strictly conserved YCGYC active site. Interestingly, no GrxC1 isoform was identified either in the genome of the two monocots (*O. sativa* and *S. bicolor*), or in EST databases from other monocot species and more generally outside dicots. On the contrary, *P. trichocarpa* and *M. truncatula* possess two GrxC1 isoforms, respectively (Table 2). It appears, from the phylogenetic tree and the identity (varying between 50.5 and 81.9%), that the GrxC1 and C2 subgroups are closely related, although the GrxC2 subgroup includes isoforms with [Y/S]CP[Y/F]C motifs, most often CPFC (Fig. 2). While there is only one member in dicots, there are two to three members in monocots, in *P. patens* and in *S. moellendorffii* (Table 2). It is tempting to speculate that the GrxC1 evolved from GrxC2 after the split between monocots and dicots. Subsequently, in a few species (*P. trichocarpa* and *M. truncatula*), the Grx C1 gene has been duplicated. From a functional point of view, it has been demonstrated that poplar and *A. thaliana* GrxC1 were unique among class I Grxs in their ability to bind a [2Fe–2S] cluster [20]. Mutagenesis experiments showed





**Fig. 3** Unrooted, NJ-based tree of class II glutaredoxins in photosynthetic organisms. The analysis was performed using MEGA 4. Branch lengths are proportional to phylogenetic distances. The name of the species is abbreviated with a *two-letter code*, except for the three *Osterococcus* species (Ost9901, OstRCC809 and Ostta). The cyanobacterial sequences form an independent group with two branches, whereas most eukaryote CGFS Grxs are clustering in the previous Grx subgroups described for higher plants, Grx S14 to S17. As for class I Grxs, a few algal members are isolated



the YCPYC active site in *P. patens* might have evolved into WCPYC in *S. moellendorffii* with the replacement of the Tyr by a Trp and then into WCSYC with the replacement of the Pro by a Ser in higher plants. The latter form apparently gave rise to WCSYS variants in some species. Interestingly, *A. thaliana* possesses the two forms, suggesting that the gene has been duplicated only in this genus, as there are also two genes in *A. lyrata* but only one in other Brassicaceae (Fig. 2, and data not shown). It has recently been demonstrated that chloroplastic poplar GrxS12 (WCSYS active site) is able to reduce substrates through a monothiol mechanism, although the protein possesses a conserved C-terminal cysteine close to the active site in the 3D structure [16]. As GrxC5 isoforms display a dithiol active site, it would be interesting to test

whether GrxC5 isoforms can function through a dithiol mechanism and have retained specific catalytic properties. The mutation of GrxS12 active site from WCSYS into YCSYS allowed the assembly of an iron sulfur center in this variant, as in GrxC1, and illustrates the importance of a single amino acid substitution in the active site of Grxs for their function [16].

The ubiquitous class II contains glutaredoxins with a CGFS active site

#### Overall view

The Grxs with a CGFS active site (later referred as CGFS Grxs) are present in most eukaryotes but not in all

prokaryotes. This analysis revealed that they are present in all photosynthetic organisms including cyanobacteria. In other prokaryotes, they are only found in proteobacteria except those of the campylobacterale order, but not in other bacterial genera nor in archaea except for the halobacteriale order. It has been suggested, from phylogenetic analyses, that it arises from a horizontal gene transfer from an ancestor of proteobacteria to halobacteriale [13]. The phylogenetic analysis of these CGFS Grxs reveals the presence (1) of six clearly identified subgroups, two in cyanobacteria and four in eukaryote organisms, which match the four subgroups (GrxS14 to S17) previously defined [8], and (2) of a few sequences isolated in the tree, as for the class I (Fig. 3). Based on an amino acid sequence comparison of 129 protein sequences from cyanobacteria and photosynthetic eukaryotes, the following characteristic motif has been identified as specific of CGFS Grxs **GX<sub>4</sub>PXCGFSX<sub>35</sub>[S/T/A]WPTXPX<sub>4</sub>GX<sub>3</sub>GG** with 13 invariant amino acids distributed along the sequence (in bold). This class contains orthologs of yeast Grx3 to 5, *E. coli* Grx4 and human Grx5 and PICOT (PKC (protein kinase C)-interacting cousin of thioredoxin) [21].

All cyanobacteria analyzed possess only one CGFS Grx isoform, whereas photosynthetic eukaryotic organisms generally have an expanded family, comprising 2–8 members distributed in four different subgroups (GrxS14 to GrxS17) (Tables 1 and 2). Most cyanobacterial CGFS Grxs cluster in two different branches which do not correspond to a single genus as, for example, Grxs from various species of *Synechococcus* are found in both subgroups. All these sequences contain a single Grx domain as the eukaryote GrxS14 and S15 subgroups, while both GrxS16 and GrxS17 types display an N-terminal extension and, in the case of GrxS17, one to three Grx domains. Because of their size and their position in the tree, the cyanobacterial Grxs might thus be more related to the eukaryote GrxS14 subgroup (Fig. 3). Nevertheless, three sequences from cyanobacteria, although similar in size, are isolated in the tree and are more related to the GrxS15 subgroup (Fig. 3). Grxs from diatoms, the red alga and the haptophyte do not cluster well (especially GrxS14 and S15 isoforms) with other eukaryote subgroups, while Grxs from green algae do, although forming separate clads. Algae and higher plants contain between 2 and 6 CGFS Grx members, whereas *P. patens* and *S. moellendorffii*, which are at an intermediate position, present an expanded class with 8 and 9 members, respectively. Indeed, *P. patens* has two GrxS14 and GrxS15 and three GrxS17 isoforms, while *S. moellendorffii* has two GrxS17 and three GrxS14 and GrxS15 isoforms. The different distribution between these two latter species suggests that the genes have been specifically duplicated in these species rather than lost in angiosperms (Table 2).

### Higher plant subgroups

GrxS14 isoforms are small proteins with a single repeat of the Grx domain and are retrieved from all analyzed genomes (Table 2). Eukaryote sequences contain an N-terminal chloroplastic targeting sequence contributing to the larger size of these isoforms (around 170 amino acids) compared to cyanobacterial isoforms (around 110 amino acids). Most organisms possess one GrxS14 isoform, but *V. vinifera*, *M. truncatula*, *P. patens* and *M. pusilla* have two and *S. moellendorffii* three (Table 2). As this property is not associated to a specific phylum or genus, the most plausible hypothesis is that a duplication occurred in isolated species.

GrxS15 isoforms constitute an independent subgroup of class II. These are also small proteins with a single repeat of the Grx domain, which display (in the case of eukaryote sequences) or not (in cyanobacterial sequences) an N-terminal mitochondrial targeting sequence. Thus, the size of these proteins oscillates between 98 and 188 amino acids in their non-maturated form. All algae have only one GrxS15 isoform. Surprisingly, there is a disparity between land plants. While the two monocots and *P. patens* or *S. moellendorffii* have two or three GrxS15 isoforms respectively, all dicots have only one (Table 2). The analysis of EST databases has confirmed this observation, as we found two different GrxS15 isoforms in two other monocot species (*Zea mays* and *Panicum virgatum*). There are several possible explanations, but the most simple is that a duplication occurred after the split with green algae, but that one gene has been lost in the dicot lineage.

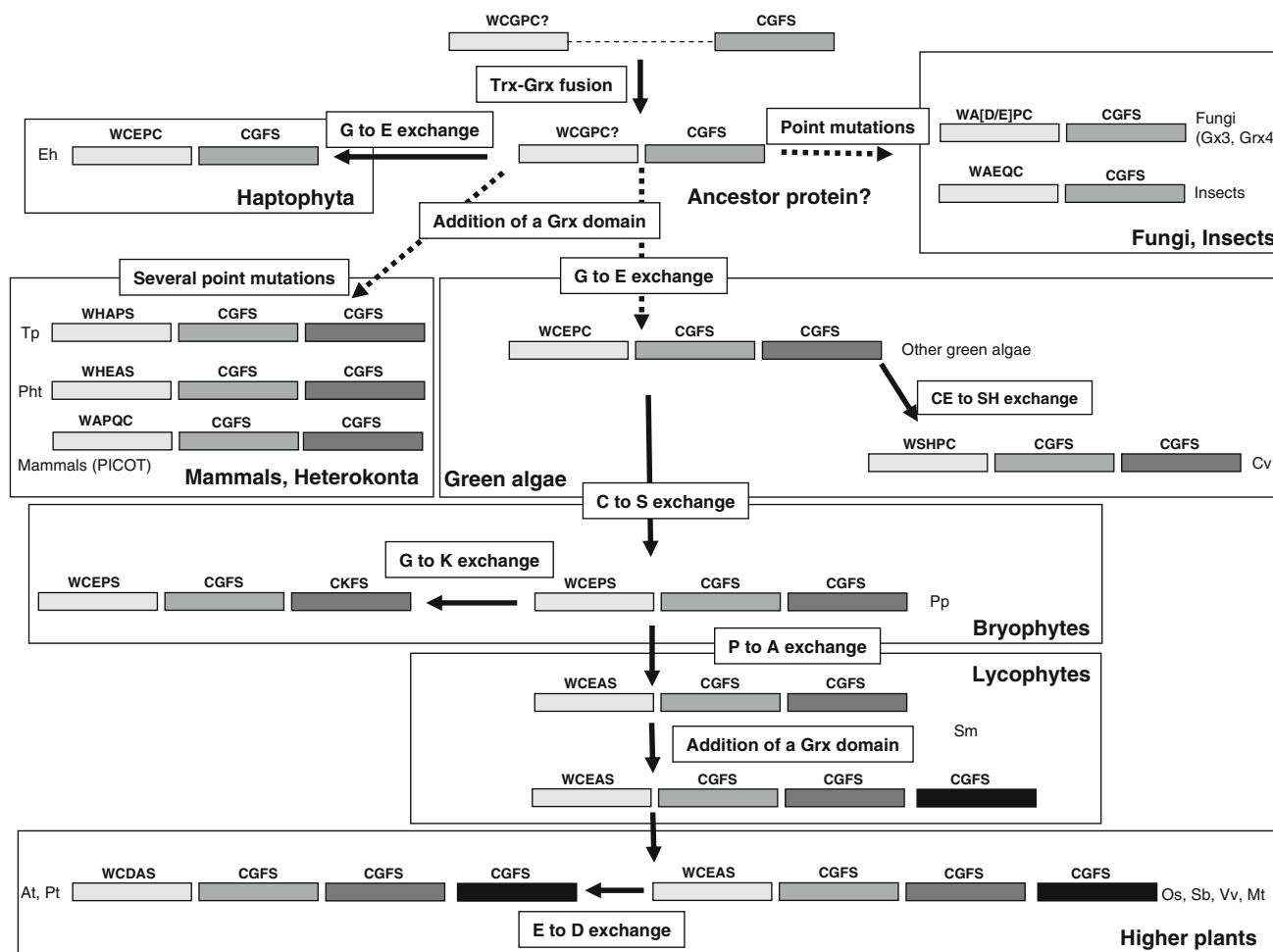
A third subgroup is constituted by GrxS16 isoforms, which are larger proteins containing, in addition to their chloroplastic targeting sequence, an N-terminal extension of unknown function. The size of these proteins is around 300 amino acids. Except for the three chromoalveolata (the two diatoms and the haptophyte) and the red alga, most other organisms possess one member (*S. bicolor* has two), suggesting that the gene is specific of the green lineage (Table 2). Interestingly, genes coding for sequences orthologous to the N-terminal domain have only been found in cyanobacteria but not at the proximity of the genes encoding the CGFS Grxs (data not shown). Thus, we can hypothesize that GrxS16 originates from the fusion of these two genes in an ancestor of green algae.

The last subgroup is constituted by larger isoforms named GrxS17. They are characterized by the presence of an N-terminal Trx-like domain linked to one, two or three Grx domains. Except for *C. merolae*, all organisms analyzed possess at least one GrxS17 isoform, suggesting that either the gene has been lost in this species or that it belongs to non-sequenced parts of this genome. This gene has been most likely duplicated in *P. patens* (3 members)

and in *P. trichocarpa* or *S. moellendorffii* (2 members) (Table 2). Some evolutionary features concerning the arrangement of GrxS17 domains merit development. It is worth mentioning that higher plants specifically possess a protein with three Grx domains, whereas other photosynthetic eukaryotes analyzed display one or two Grx domains, and other eukaryotes, such as fungi (Grx3 and Grx4) or mammals (PICOT protein), one or two Grx domains (Fig. 4) [21, 22]. This type of fusion protein is absent in prokaryotes.

We have tentatively reconstituted the evolution of the GrxS17 sequences from two initial Trx and CGFS Grx encoding genes (Fig. 4). It is likely that the initial event is the fusion of the two genes in a common ancestor to most eukaryotes (Fig. 4). While the domain arrangement was not modified in fungi and haptophyta, algae and mammals

acquired a second Grx domain, which most likely originated from a duplication of the Grx domain only. Indeed, it is believed that these repeats are created by internal duplication, where the duplicated domain is inserted in frame after the original domain [23]. Finally, another similar event could have led to the acquisition of a third Grx module in lycopphyte, as *S. moellendorffii* contains two different forms with two or three Grx domains. As shown in Fig. 4, whereas the Grx active sites were essentially unaltered (we identified only one sequence in *P. patens* with a CGKS instead of the CGFS motif), the thioredoxin active site diverged between all species by the introduction of point mutations. The usual WCGPC Trx active site may have been modified through mutations to give the present sequences of haptophyta, fungi and mammalian orthologs (respectively *E. huxleyi* GrxS17, fungal Grx3 or Grx4 and



**Fig. 4** Schematic representation of the putative reconstituted evolution of GrxS17 orthologs in living organisms. Trx and Grx domains are represented by rectangles with the active site indicated on top. The putative major steps for the formation of higher plants GrxS17 isoforms from cyanobacterial members may consist in three consecutive events (shaded boxes): (1) Trx-Grx gene fusion; (2) addition of a Grx domain in a common ancestor of heterokonta and green algae;

and (3) addition of a second Grx domain in lycopphytes. As the whole fusion has not been duplicated entirely, we hypothesize that the addition of a Grx domain in the C-terminal part of the protein is arising from a partial duplication. We have also indicated the possible evolution of the active site in the thioredoxin domain by indicating the point mutations required (open boxes)

mammalian PICOT). Following the addition of a Grx domain, the active site of the Trx domain has been transformed in heterokonta into WHAPS and WHEAS. In most green algae, except *Chlorella vulgaris*, a single mutation transformed the Trx active site into WCEPC and this is most likely the starting point for subsequent modifications/mutations in the green lineage. The most plausible scenario is that this motif has been modified into WCEPS in bryophytes and subsequently into WCEAS in lycophytes. In higher plants, some proteins conserved the WCEAS sequence (*O. sativa*, *S. bicolor*, *M. truncatula*, *V. vinifera*), while others evolved into a WCDAS sequence (*P. trichocarpa*, *A. thaliana*). All these modifications suggest that the Trx module is no longer active, but it could be useful for folding or interacting with other proteins.

Together with proteins of class I, the Grx class II is also widespread and conserved in photosynthetic organisms. The differences in primary sequences, in domain organisation and in sub-cellular localisation, suggest that Grxs of this class should have distinct functions from class I Grxs and between each other. To date, only a few CGFS Grxs from photosynthetic organisms have been studied. From these studies, it appears that most CGFS Grxs from cyanobacteria, algae and higher plants can bind a labile iron sulfur cluster (ISC) of the [2Fe–2S] type [24, 25]. This is consistent with the mutagenesis data generated with GrxC1, showing that the presence of the glycine after the catalytic cysteine is crucial for ISC incorporation [20]. In vitro, plant GrxS14 is able to transfer rapidly and stoichiometrically this center intact to a chloroplastic apoferritin, providing evidence for a role of CGFS Grxs as [2Fe–2S] cluster donors for the maturation of chloroplastic Fe–S proteins [24]. This role is supported by the fact that GrxS14, S16 and S17, but not GrxS15, can complement the defect in mitochondrial iron-sulfur cluster biogenesis of a yeast *grx5* strain [24]. The fact that the cluster is labile suggests that these proteins can have dual functions. Indeed, it has been demonstrated that the apoform of CrGrx3 is able to perform in vitro the deglutathionylation of a glutathionylated A<sub>4</sub>-GAPDH with an atypical reaction mechanism. During the catalytic process, CrGrx3 forms an intramolecular disulfide bridge, with a quite low redox potential, which is not reduced by GSH but by ferredoxin-thioredoxin reductase [5].

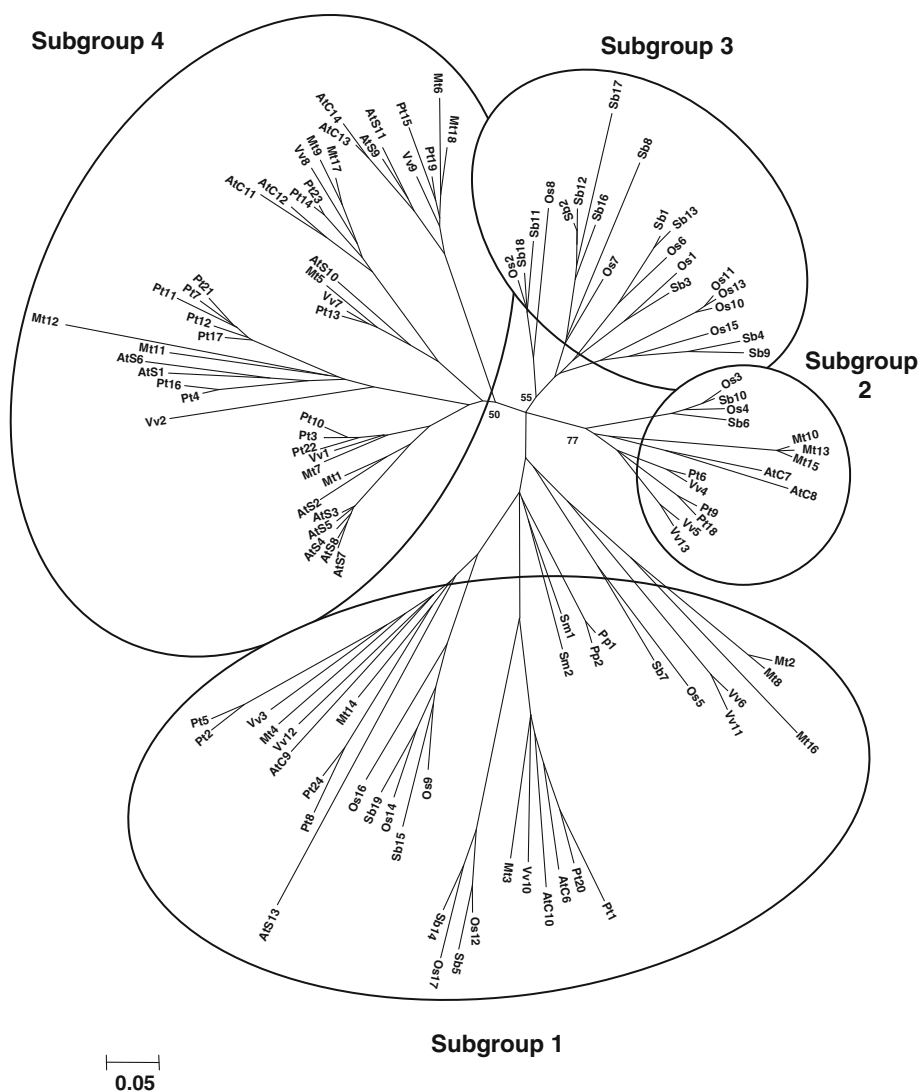
The glutaredoxins of class III are specific to terrestrial plants

Subclass III corresponds generally to Grxs with a CCxx active site and a general C[C/Y/G/S/P/F][M/L/F][C/S/G/I/A/T]X<sub>39</sub>[S/T/A/P/N/V/Q/K/L/G/R][V/L/A/S/P/F/I]**PX<sub>9</sub>G**[G/S/A/N/P/T] motif with three invariant amino acids (in bold). As there are no orthologous sequences in cyanobacteria and algae,

this analysis reveals that class III is specific to terrestrial plants, but with a large differential distribution between higher and lower plants. Indeed, *P. patens* and *S. moellendorffii* contains only two genes, whereas the number of isoforms in higher plants ranges from 13 in *V. vinifera* to 24 in *P. trichocarpa* (Table 2). These observations suggest that genes encoding CCxx Grx isoforms appeared late in the green lineage evolution and that they have been subjected to many duplication events in higher plants. This is exemplified both in *A. thaliana* and in *P. trichocarpa*, where five isoforms of subgroup 4 (GrxS3, S4, S5, S7 and S8 for *A. thaliana* and Grx Pt7, Pt11, Pt12, Pt17 and Pt21 for *P. trichocarpa*) are clustering together and exhibit a very high identity (from 91 to 95% for *A. thaliana* isoforms and from 84 to 90% for *P. trichocarpa* isoforms), suggesting a recent duplication. In addition, in *A. thaliana*, the five genes occupy adjacent positions in the genome. Nevertheless, compared to Grxs from classes I and II, which are all expressed, the absence of ESTs for many class III Grx genes could indicate that not all of them are expressed and functional. As only a few isoforms have been characterized, future data will be needed to understand the roles of CCxx Grx isoforms in plant physiology and metabolism and why higher plants possess so many isoforms compared to mosses and lycophytes.

Phylogenetic analyses demonstrate that class III Grxs can be subdivided into four major distinct subgroups (Fig. 5). Thanks to the multiplicity of sequences used in this analysis, the subgroups are well delineated. It is interesting to note that the difference is not based on the presence of mono- or di-thiol active sites as, for example, Grxs with a CCMC active site motif belong to all four subgroups. On the contrary, subgroups 3 and 4 are rather specific respectively to monocots and dicots, while subgroups 1 and 2 contain dicot and monocot Grx isoforms (Fig. 5). Subgroup 1 comprises Grx isoforms from all land plants and particularly Grxs from *P. patens* and *S. moellendorffii*, which are likely to be the ancestral forms and represent one branch (Fig. 5). This subgroup is quite heterogeneous, since there are three other major branches with both monocot and dicot sequences. In addition, it contains most Grx isoforms with atypical active site motifs, either from dicots (AtGrxS13, VvGrx14 and MtGrx16 possess CCLG, CCMT and CCFS active sites) or from monocots (OsGrx17, SbGrx14 and OsGrx14 possess CYMA, CCMA and CCLI active sites) (ESM, sequence file). In this subgroup, the only Grx member functionally characterized (AtGrxC9) is involved in pathogen response [26]. Subgroup 2 is the smallest subgroup, but it is well delineated and it contains the CCxx Grx isoforms best characterized to date. Different studies, dealing with AtGrxC7 and AtGrxC8, also called ROXY1 and ROXY2, have demonstrated that these Grxs are needed for the development

**Fig. 5** Unrooted, NJ-based tree of class III glutaredoxins from land plants. The analysis was performed using MEGA 4. Branch lengths are proportional to phylogenetic distances. The name of the protein is composed of the *two-letter code* for the species followed by the *number* of the isoform, for example, 1–24 for *Populus*. The exception is for *A. thaliana* sequences which follow the nomenclature defined in [2], with a *C* or an *S* indicating the dithiol or monothiol nature of the active site. The proteins are distributed into four subgroups (see the text for their description). The sequences from cyanobacteria form an independent group with two branches, whereas most eukaryote CGFS Grxs are clustering in the previous Grx subgroups described for higher plants, Grx S14 to S17. As for class I Grxs, a few algal members are isolated



of reproductive tissues [12, 27]. Recently, ROXY1 was shown to be localized in the nucleus and to interact with TGA transcription factors, an interaction required for correct petal development in *A. thaliana* [28]. Since they are selectively distributed in the phylogenetic tree, orthologs from other species could have a similar function.

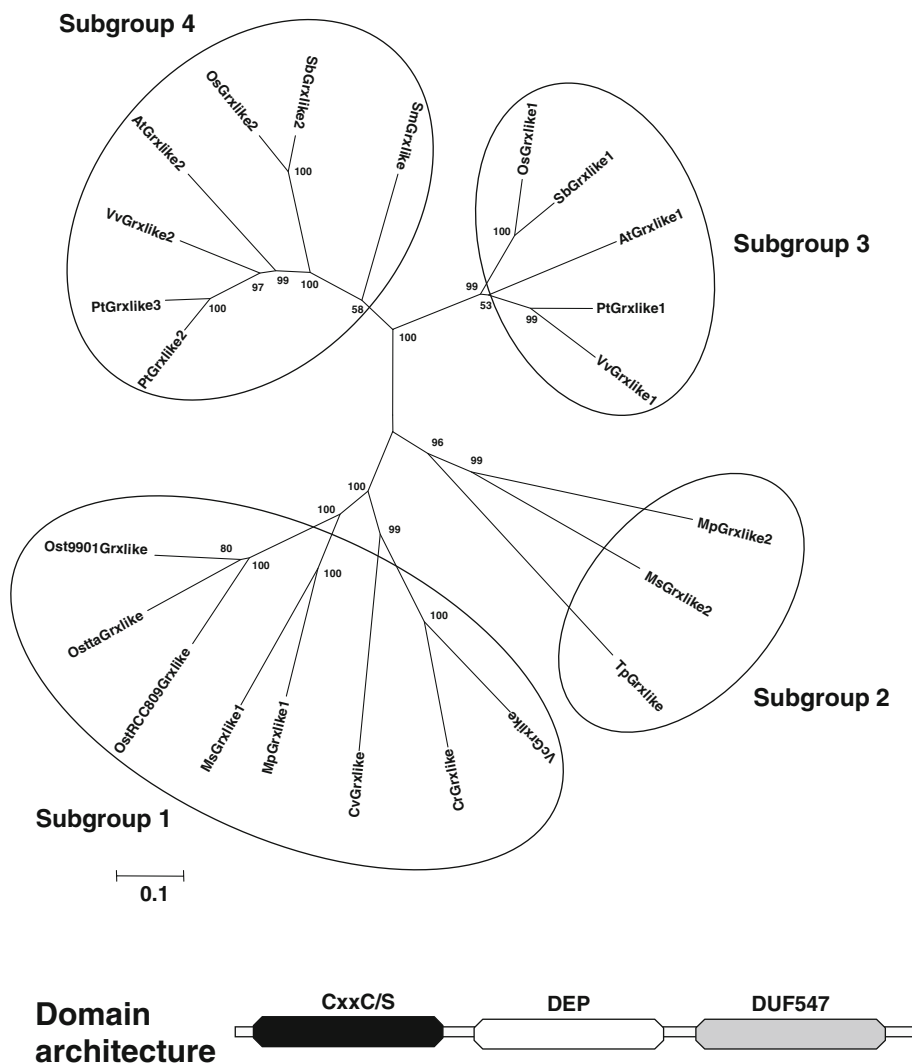
Subgroup 3 is only constituted by monocot Grxs, which cluster into four branches (always containing both *S. bicolor* and *O. sativa* Grxs). These branches become much more defined when we add sequences from other monocots (data not shown). Subgroup 4 is restricted, as mentioned before, to dicots. Similarly to subgroup 3, there are four branches, which contain at least one member of each species, but they are better delineated than in subgroup 3 because a higher number of sequences have been used. The differences in the number of members in a given species most likely arise from duplication.

Fusion proteins between a Grx and two additional modules in eukaryotes constitute the new class IV glutaredoxins

It was previously considered that, in land plants, the Grx family could be subdivided into three classes [2, 7, 8]. Our current analyses have revealed the existence of one additional Grx class in all eukaryotes (class IV). These Grxs exhibit an N-terminal Grx domain with all the characteristic motifs, fused to two other domains (DEP and DUF547 based on pfam conserved domains) of unknown function in the C-terminal part and have been called Grx-like (Fig. 6). Not much information is available for these two domains. A study in yeast demonstrated that the signaling protein Sst2 is interacting specifically with G protein-coupled receptors, with the DEP domain of Sst2 generating the flexibility required for protein–protein interaction [29].



**Fig. 6** Unrooted, NJ-based tree of class IV glutaredoxins in photosynthetic eukaryotes. The analysis was performed using MEGA 4. Branch lengths are proportional to phylogenetic distances. Proteins of this class are clustering into four subgroups, two for algal members and two for terrestrial plant members. The domain architecture of these Grxs is represented in the *lower part* of the figure



Depending on the organisms, the size of the proteins is very variable and ranges from 550 to 730 amino acids for algal and higher plant sequences respectively. The difference is due to the presence of an N-terminal extension situated before the Grx domain, but it is not predicted to contain a signal sequence. The Grx active site sequence is also quite heterogeneous, with a majority of CxDC/S motif in higher plants and of CPxC motif in green algae. In several respects (active site sequence, amino acid involved in glutathione binding), the Grx domain from algal Grx-like sequences resembles Grxs from class I, while land plants' Grx-like sequences diverged quite a lot. The Grx-domain is characterized by the presence of the motif **C**[P/R]<sub>S/Q/E</sub>[D/H/Y/E/F]**C**/<sub>S</sub>X<sub>36</sub>[S/T/Q/R/A]**V**/<sub>A</sub>**P**X<sub>9</sub>**G**[G/S] with three invariant amino acids (in bold) (ESM, Fig. 2).

At least one Grx sequence was retrieved from all analyzed genomes except from *E. huxleyi*, *P. tricornutum*, *C. merolae* and more surprisingly *P. patens* genomes, but we cannot completely rule out that these organisms do not

possess a type IV Grx isoform (Fig. 6 and Table 2). Moreover, as cyanobacteria do not possess this protein, it has probably appeared early in the green lineage evolution, but after the split with cyanobacteria. It is worth mentioning that the number of Grx-like isoforms increased during evolution since the green algae, *T. pseudonana* and *S. moellendorffii* have only one isoform, whereas higher plants possess two isoforms, except poplar which has three isoforms (Table 2). Hence, a duplication most likely occurred in an ancestor organism of flowering plants after the split with lycopphyte and *P. patens* might have lost the gene. An additional duplication occurred in poplar.

Phylogenetic analyses show that these proteins can be classified into four subgroups (Fig. 6). Subgroup 1 only includes Grx isoforms from green algae with a CP[H/Y]C active site, whereas subgroup 2 corresponds to isoforms of *T. pseudonana* and *Micromonas*, which possess a C[S/R][H/F]C active site. The two last subgroups contain terrestrial plant members. Both subgroups 3 and 4 contain

monocot and dicot isoforms, but sequences from subgroup 3 display a CRD[C/S] active site, while sequences from subgroup 4, including the *S. moellendorffii* member, exhibit a C[P/E/Q][D/E][C/S] motif (Fig. 6). Orthologs were also retrieved in a limited number of animal genomes such as *Salmo salar*, *Dario rerio*, *Ciona intestinalis*, *Trichoplax adhaerens*, *Tetraodon nigroviridis*, but not from bacterial, mammalian and fungal genomes, suggesting that these proteins are restricted to specific eukaryote organisms (data not shown). To date, there are no experimental data concerning the functions of these proteins.

Bimodular fusion proteins constitute two Grx classes (V and VI) specific to cyanobacteria

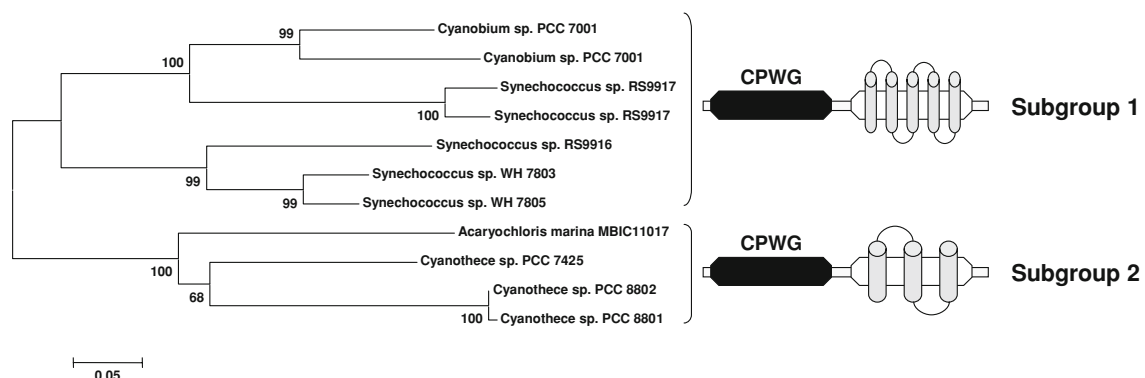
As observed in photosynthetic eukaryotes, cyanobacterial Grx isoforms can be classified into four different classes, but two are newly identified classes (class V and VI), containing elongated Grxs, and unique to cyanobacteria. The class V contains proteins with two domains, (1) an N-terminal Grx module with an atypical CPWG active site, a general CPWGX<sub>35</sub>[T/S]TPX<sub>9</sub>GG pattern, and (2) a C-terminal module containing, from the TMHMM prediction tool, 3 or 5 transmembrane domains of different sizes (Fig. 7). These alternative arrangements are consistent with the existence of two major subgroups as determined by phylogenetic analyses. The resulting hybrid protein could be present in some chroococcalles and in proteobacteria (data not shown). The C-terminal module is present as an isolated protein of 150 amino acids in a few *Prochlorococcus* species, which do not possess the fusion protein (data not shown). The sequences from proteobacteria possess a slightly modified active site, generally CPYG. The size of the proteins ranges from 230 to 280 amino acids. In subgroups 1 and 2, the identity is comprised between 55.6 and 92.4% or between 63.9 and 99.6%, respectively, and it ranges from 44.8 to 56.4% between members of the two

subgroups. Because of the few representatives found, there are many conserved residues in both domains (30 amino acids conserved in the Grx domain of the 10 sequences identified) (ESM, Fig. 3). Two cysteine residues in a CAC motif are conserved in the C-terminal domain. The topology predictions indicate that they are situated in a loop between two transmembrane domains and located at the same side from the Grx domain, suggesting that a redox control could occur between the two domains.

Class VI also includes proteins with two domains, an N-terminal domain of unknown function (DUF296 superfamily) followed by a Grx domain with an atypical CPW[C/S] active site and more generally a CP[W/F][C/S]X<sub>35</sub>[T/S/V][F/V]PX<sub>9</sub>G[G/D] motif (Fig. 8). Among the 21 sequences, there are seven conserved residues in the Grx domain (ESM, Fig. 4). The size of the proteins is around 200 amino acids and the identity is between 32.5 and 99%. This hybrid form was detected only in chroococcalles and, more precisely, in some organisms of the *Prochlorococcus*, *Synechococcus* and *Cyanobium* genera, but neither in other cyanobacteria nor in other bacteria, indicating that it has been lost early along the green lineage evolution. A gene family coding for proteins of 130 amino acids, related to the N-terminal part of this hybrid protein, is found in some proteobacteria. The identity is around 30% between these protein sequences (data not shown). To date, the putative function of these two proteins is unknown, but the presence of two conserved cysteine residues in the N-terminal domain might indicate it is redox-regulated.

Genome analysis for gene fusion, gene clustering, and gene co-occurrence

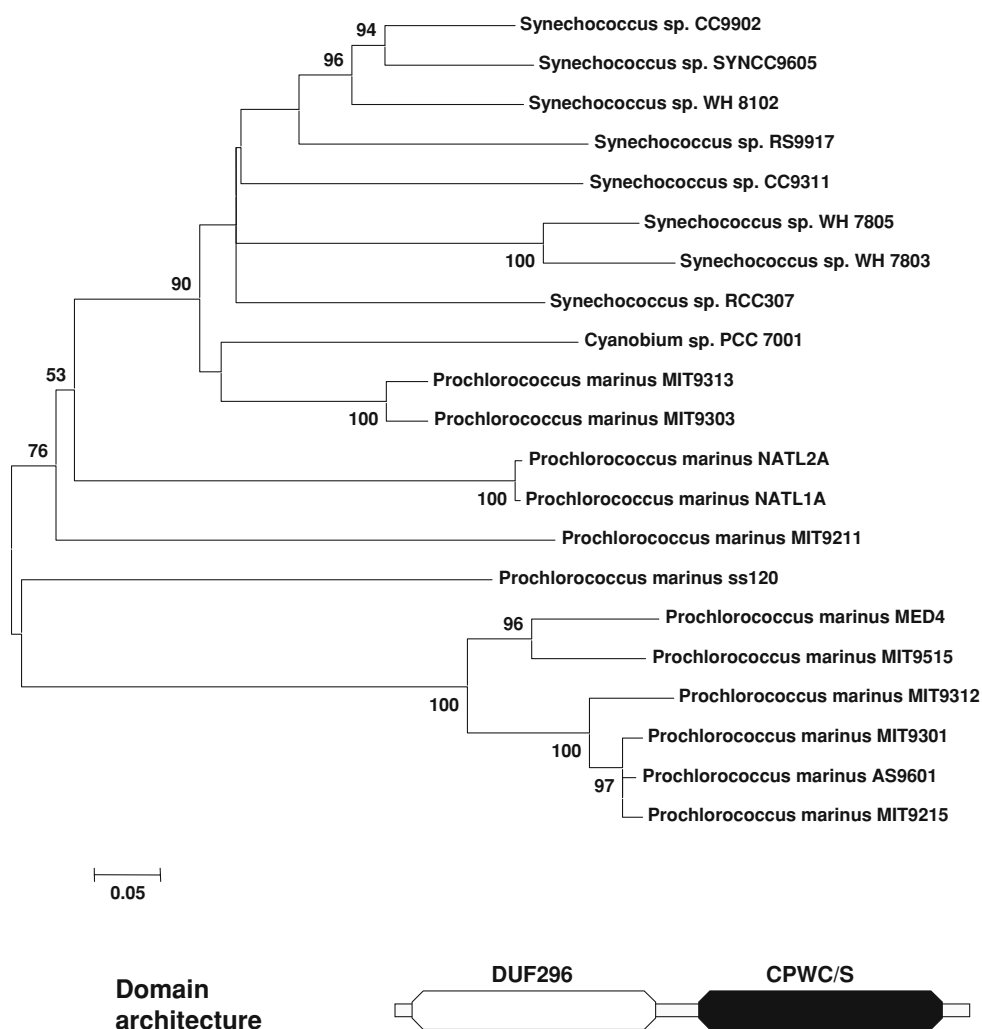
Another resource offered by genomic data is to predict functional or physical interactions between proteins by analyzing (1) the presence of fusion proteins, (2) the conservation of gene order in presumptive bacterial operons,



**Fig. 7** Unrooted, NJ-based tree of class V glutaredoxins in cyanobacteria. The analysis was performed using MEGA4. Branch lengths are proportional to phylogenetic distances. Proteins of this class are

clustering into two separate branches depending on the organisation of the C-terminal transmembrane part

**Fig. 8** Unrooted, NJ-based tree of class VI glutaredoxins in cyanobacteria. The analysis was performed using MEGA4. Branch lengths are proportional to phylogenetic distances. The domain architecture of these Grxs is represented in the *lower part* of the figure. This protein is only present in most cyanobacteria of the order Synechococales but not all, and in a closely related Chroococcales, *cyanobium* sp. PCC7001



and (3) the co-occurrence of genes in several organisms. When combined, these analyses constitute a powerful tool for identifying putative Grx partners both in prokaryotes and eukaryotes (where the gene order is not as important for their regulation), provided that there are orthologous sequences.

For example, it has been demonstrated, in parallel studies in 2001, that a class of plant peroxiredoxin (Prx), called type II Prx, was able to reduce various peroxides with electrons supplied by a plant Grx assisted by a glutathione regeneration system and that a PrxGrx hybrid protein from *Chromatium gracile* is efficiently regenerated using an analog of glutathione, called glutathione amide [30, 31]. It was confirmed later that several PrxGrx from pathogenic bacteria and cyanobacteria are reduced by glutathione [18, 19]. This is a remarkable example of how the analysis of bacterial genomes and sequences could predict interactions or collaborations between proteins in eukaryotes. There are many other fusion proteins in non-

photosynthetic organisms containing a Grx domain and hence many new putative partners [13].

Through the identification of putative operons containing *grx* genes in cyanobacteria, we aimed at identifying new putative target proteins of Grxs. From the “protein clusters” tool available at the NCBI webpage, we have looked at genes present on both sides of *grx* genes, regardless of their orientation. We have selected three examples, for three different Grx classes, where the gene order is highly conserved and where it is largely distributed among cyanobacteria and/or bacteria.

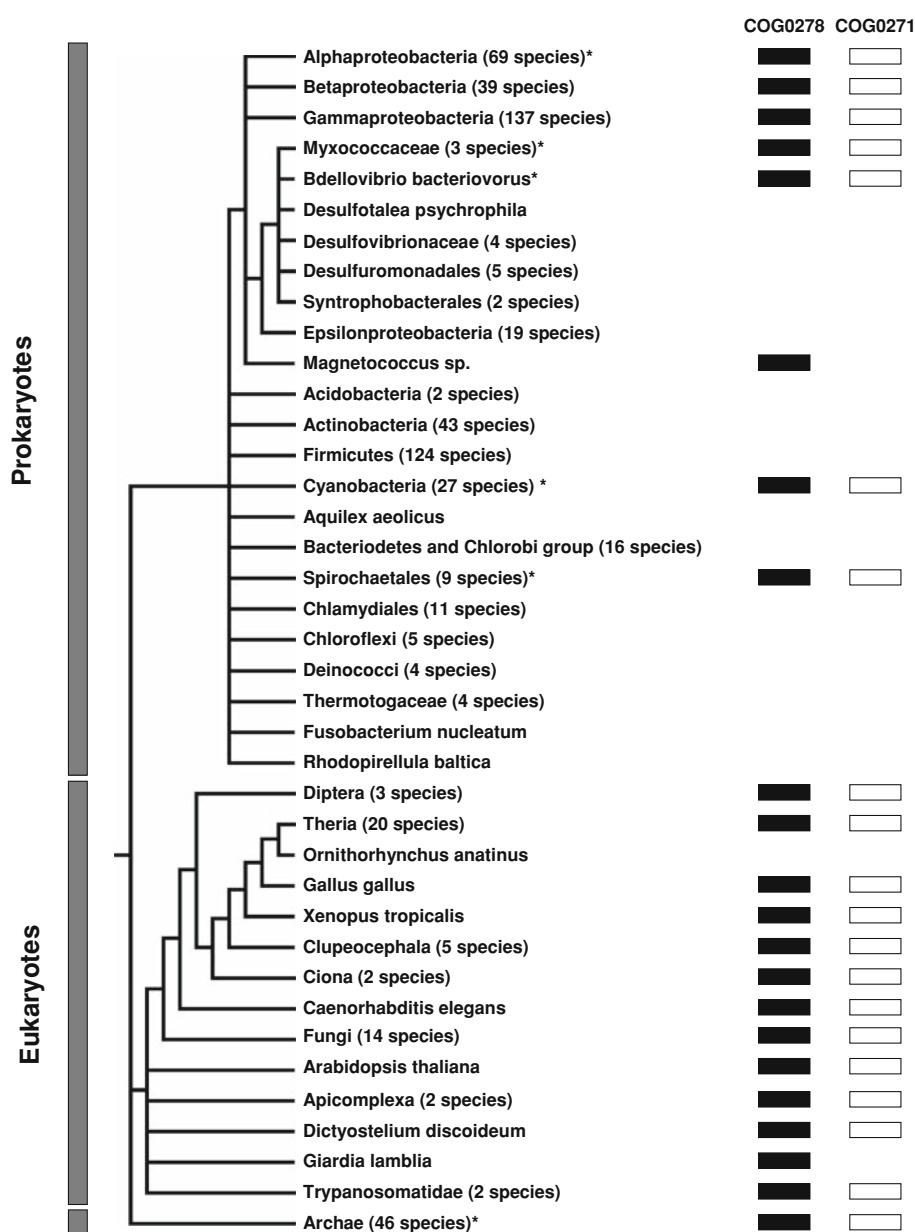
In Chroococcales and Prochlorococcales, but not in the available sequenced Nostocales, the class I *grx* gene is very frequently associated to the glutathione synthetase encoding gene (*gshB*) and in the same orientation. Interestingly, the two proteins derived from these two genes are physiologically linked, as class I Grx isoforms require glutathione for their functioning. Alternatively, because of their disulfide reductase activity, Grxs might regulate the

activity of glutathione synthetases at the post-translational level. In the amino acid sequence comparison of cyanobacterial glutathione synthetases, there is one conserved cysteine in the C-terminal part which could be subjected to S-thiolation. This organisation is typical of cyanobacteria as it is not found in other bacteria. In addition, this analysis does not reveal a large genome co-occurrence, as some organisms possess either only a *grx* gene or a *gshB* gene or none of them. This observation suggests several comments: (1) the GSH/Grx system is not ubiquitous, (2) in organisms lacking glutathione, but possessing a Grx, these Grxs do not necessarily use glutathione and might be reduced by thioredoxin reductases (this has already been proven in some organisms), and (3) in organisms lacking Grxs, but

possessing glutathione, glutathione is involved in Grx-independent metabolic pathways.

Cyanobacteria and most bacteria have only one CGFS Grx, constituted by a single Grx domain. In cyanobacteria, most CGFS Grx encoding genes are adjacent to a BolA or YrbA encoding gene (mostly in the 5' part but there are a few examples where the BolA gene is downstream), and both genes are present in the same transcriptional orientation, suggesting that both proteins interact in vivo. This gene clustering is also found in many other prokaryote genomes, essentially in  $\alpha$ -proteobacteria and some other bacterial classes, but not in  $\beta$ - and  $\gamma$ -proteobacteria (Fig. 9). On the contrary, the genes surrounding these two genes are more variable. This observation is also supported

**Fig. 9** Gene clustering and co-occurrence between CGFS Grxs and BolA. Analysis of gene co-occurrence between CGFS Grx (COG0278) and BolA (COG0271) genes was performed using the “STRING” database (<http://string.embl.de/>) [15]. From 630 organisms analyzed, only two (*Magnetococcus* sp. and *Giardia lamblia*) possess only one of the two genes. Moreover, when CGFS Grx and BolA genes are clustering in all or only some prokaryote organisms of a given phylum, an *asterisk* has been added. The two genes are clustering in the large cyanobacteria and alpha-proteobacteria genus, whereas, except in a few cases, they do not cluster in beta- and gamma-proteobacteria. The analysis of the clustering of Grx genes was performed using “protein clusters” tool available at the NCBI webpage and the Microbial Genome Database (MGDB, <http://mbgd.genome.ad.jp/>)



by the strong genome co-occurrence which exists for these two genes. Indeed, with a few exceptions (for example, *Magnetococcus* sp. and *Giardia lamblia*), all organisms having CGFS Grxs also possess a BolA member and the opposite is true, all organisms lacking CGFS Grxs do not possess a BolA member (Fig. 9). From the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database which enables the analysis of 630 organisms, we detected 293 organisms possessing both CGFS Grx and BolA genes, whereas 298 organisms do not contain CGFS Grx and BolA homologs (Fig. 9) [15]. The function of BolA, small proteins of around 90 amino acids, is not absolutely clear, but they are defined as transcriptional regulators. Huynen et al. previously suggested from a bioinformatic analysis that BolA would interact with CGFS Grxs without mentioning precisely how [32]. It is only recently, that, in yeast, an *in vivo* complex between Grx3 and Grx4 (two CGFS Grxs with an N-terminal Trx-like module), a BolA protein (Fra2: Fe repressor of activation-2) and an aminopeptidase P-like protein (Fra1: Fe repressor of activation-1) has been isolated [33].

From the genome analysis of bacteria possessing class V Grxs, we found that this *grx* gene is frequently associated to genes encoding proteins related to the MerR transcriptional factor. The two genes are orientated in an opposite direction, but it is known that oppositely oriented genes, especially pairs of genes with one transcriptional regulator, may be functionally linked [34]. In some cases, the two genes are separated by one or more genes, but still in the same genome region. This family of metal sensing regulators, initially identified as mercuric response factor, is composed of proteins which bind to several metals and subsequently regulate the transcription of metal-specific genes [35]. In cyanobacteria, the MerR identified are small proteins of 140 amino acids with two conserved cysteines in the C-terminal part in a CxxxC motif (data not shown), whereas in  $\alpha$ -proteobacteria, the proteins have approximately the same size, but the two conserved cysteines, also included in a CxxxC motif, are not present at the same position (data not shown), suggesting that they belong to different MerR subgroups.

## Concluding remarks

The present phylogenetic analysis confirms that Grxs constitute a very complex and diversified group in all photosynthetic prokaryotes and eukaryotes, but the Grx classes and the gene contents are different. Grx isoforms of class I and II remained relatively conserved between prokaryotes and eukaryotes during the evolution process, suggesting that they fulfil fundamental functions for cellular life. On the other hand, some newly identified Grx

classes are restricted either to prokaryotes or to eukaryotes, suggesting that they are involved in processes specific to prokaryotic or eukaryotic organisms. Future studies will be needed to confirm that these proteins are active as glutaredoxins and to understand their physiological role. In addition, the identification, through analysis of fusion proteins, gene clustering and co-occurrence, of putative Grx partners, such as BolA, will have to be experimentally confirmed.

## References

1. Fernandes AP, Fladvad M, Berndt C, Andresen C, Lillig CH, Neubauer P, Sunnerhagen M, Holmgren A, Vlamis-Gardikas A (2005) A novel monothiol glutaredoxin (Grx4) from *Escherichia coli* can serve as a substrate for thioredoxin reductase. *J Biol Chem* 280:24544–24552
2. Rouhier N, Gelhaye E, Jacquot JP (2004) Plant glutaredoxins: still mysterious reducing systems. *Cell Mol Life Sci* 61:1266–1277
3. Johansson C, Lillig CH, Holmgren A (2004) Human mitochondrial glutaredoxin reduces S-glutathionylated proteins with high affinity accepting electrons from either glutathione or thioredoxin reductase. *J Biol Chem* 279:7537–7543
4. Reynolds CM, Meyer J, Poole LB (2002) An NADH-dependent bacterial thioredoxin reductase-like protein in conjunction with a glutaredoxin homologue form a unique peroxiredoxin (AhpC) reducing system in *Clostridium pasteurianum*. *Biochemistry* 41:1990–2001
5. Zaffagnini M, Michelet L, Massot V, Trost P, Lemaire SD (2008) Biochemical characterization of glutaredoxins from *Chlamydomonas reinhardtii* reveals the unique properties of a chloroplastic CGFS-type glutaredoxin. *J Biol Chem* 283:8868–8876
6. Rodriguez-Manzanique MT, Ros J, Cabisco E, Sorribas A, Herrero E (1999) Grx5 glutaredoxin plays a central role in protection against protein oxidative damage in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19:8180–8190
7. Lemaire SD (2004) The glutaredoxin family in oxygenic photosynthetic organisms. *Photosynth Res* 79:305–318
8. Rouhier N, Couturier J, Jacquot JP (2006) Genome-wide analysis of plant glutaredoxin systems. *J Exp Bot* 57:1685–1696
9. Tamarit J, Belli G, Cabisco E, Herrero E, Ros J (2003) Biochemical characterization of yeast mitochondrial Grx5 monothiol glutaredoxin. *J Biol Chem* 278:25745–25751
10. Meyer Y, Siala W, Bashandy T, Riondet C, Vignols F, Reichheld JP (2008) Glutaredoxins and thioredoxins in plants. *Biochim Biophys Acta* 1783:589–600
11. Rouhier N, Lemaire SD, Jacquot JP (2008) The role of glutathione in photosynthetic organisms: emerging functions for glutaredoxins and glutathionylation. *Annu Rev Plant Biol* 59:143–166
12. Xing S, Lauri A, Zachgo S (2006) Redox regulation and flower development: a novel function for glutaredoxins. *Plant Biol (Stuttg)* 8:547–555
13. Alves R, Vilaprinyo E, Sorribas A, Herrero E (2009) Evolution based on domain combinations: the case of glutaredoxins. *BMC Evol Biol* 9:66
14. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599
15. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, Doerks T, Julien P, Roth A, Simonovic M, Bork P, von Mering C



- (2009) STRING 8 – a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* 37:D412–D416
16. Couturier J, Koh CS, Zaffagnini M, Winger AM, Gualberto JM, Corbier C, Decottignies P, Jacquot JP, Lemaire SD, Didierjean C, Rouhier N (2009) Structure-function relationship of the chloroplastic glutaredoxin S12 with an atypical WCSYS active site. *J Biol Chem* 284:9299–9310
  17. Pauwels F, Vergauwen B, Vanrobaeys F, Devreese B, Van Beeumen JJ (2003) Purification and characterization of a chimeric enzyme from *Haemophilus influenzae* Rd that exhibits glutathione-dependent peroxidase activity. *J Biol Chem* 278:16658–16666
  18. Rouhier N, Jacquot JP (2003) Molecular and catalytic properties of a peroxiredoxin-glutaredoxin hybrid from *Neisseria meningitidis*. *FEBS Lett* 554:149–153
  19. Hong SK, Cha MK, Kim IH (2008) A glutaredoxin-fused thiol peroxidase acts as an important player in hydrogen peroxide detoxification in late-phased growth of *Anabaena* sp. PCC7120. *Arch Biochem Biophys* 475:42–49
  20. Rouhier N, Unno H, Bandyopadhyay S, Masip L, Kim SK, Hirasawa M, Gualberto JM, Lattard V, Kusunoki M, Knaff DB, Georgiou G, Hase T, Johnson MK, Jacquot JP (2007) Functional, structural, and spectroscopic characterization of a glutathione-ligated [2Fe-2S] cluster in poplar glutaredoxin C1. *Proc Natl Acad Sci USA* 104:7379–7384
  21. Herrero E, de la Torre-Ruiz MA (2007) Monothiol glutaredoxins: a common domain for multiple functions. *Cell Mol Life Sci* 64:1518–1530
  22. Morel M, Kohler A, Martin F, Gelhaye E, Rouhier N (2008) Comparison of the thiol-dependent antioxidant systems in the ectomycorrhizal *Laccaria bicolor* and the saprotrophic *Phanerochaete chrysosporium*. *New Phytol* 180:391–407
  23. Bjorklund AK, Ekman D, Elofsson A (2006) Expansion of protein domain repeats. *PLoS Comput Biol* 2:e114
  24. Bandyopadhyay S, Gama F, Molina-Navarro MM, Gualberto JM, Claxton R, Naik SG, Huynh BH, Herrero E, Jacquot JP, Johnson MK, Rouhier N (2008) Chloroplast monothiol glutaredoxins as scaffold proteins for the assembly and delivery of [2Fe-2S] clusters. *EMBO J* 27:1122–1133
  25. Picciocchi A, Saguez C, Boussac A, Cassier-Chauvat C, Chauvat F (2007) CGFS-type monothiol glutaredoxins from the cyanobacterium *Synechocystis* PCC6803 and other evolutionary distant model organisms possess a glutathione-ligated [2Fe-2S] cluster. *Biochemistry* 46:15018–15026
  26. Ndamukong I, Abdallat AA, Thurow C, Fode B, Zander M, Weigel R, Gatz C (2007) SA-inducible *Arabidopsis* glutaredoxin interacts with TGA factors and suppresses JA-responsive PDF1.2 transcription. *Plant J* 50:128–139
  27. Xing S, Zachgo S (2008) ROXY1 and ROXY2, two *Arabidopsis* glutaredoxin genes, are required for anther development. *Plant J* 53:790–801
  28. Li S, Lauri A, Ziemann M, Busch A, Bhave M, Zachgo S (2009) Nuclear activity of ROXY1, a glutaredoxin interacting with TGA factors, is required for petal development in *Arabidopsis thaliana*. *Plant Cell* 21:429–441
  29. Ballon DR, Flanary PL, Gladue DP, Konopka JB, Dohlman HG, Thorner J (2006) DEP-domain-mediated regulation of GPCR signaling responses. *Cell* 126:1079–1093
  30. Rouhier N, Gelhaye E, Sautiere PE, Brun A, Laurent P, Tagu D, Gerard J, de Fay E, Meyer Y, Jacquot JP (2001) Isolation and characterization of a new peroxiredoxin from poplar sieve tubes that uses either glutaredoxin or thioredoxin as a proton donor. *Plant Physiol* 127:1299–1309
  31. Vergauwen B, Pauwels F, Jacquemotte F, Meyer TE, Cusanovich MA, Bartsch RG, Van Beeumen JJ (2001) Characterization of glutathione amide reductase from *Chromatium gracile*. Identification of a novel thiol peroxidase (Prx/Grx) fueled by glutathione amide redox cycling. *J Biol Chem* 276:20890–20897
  32. Huynen MA, Spronk CA, Gabaldon T, Snel B (2005) Combining data from genomes, Y2H and 3D structure indicates that BolA is a reductase interacting with a glutaredoxin. *FEBS Lett* 579:591–596
  33. Kumanovics A, Chen OS, Li L, Bagley D, Adkins EM, Lin H, Dingra NN, Outten CE, Keller G, Winge D, Ward DM, Kaplan J (2008) Identification of FRA1 and FRA2 as genes involved in regulating the yeast iron regulon in response to decreased mitochondrial iron-sulfur cluster synthesis. *J Biol Chem* 283:10276–10286
  34. Korb J, Jensen LJ, von Mering C, Bork P (2004) Analysis of genomic context: prediction of functional associations from conserved bidirectionally transcribed gene pairs. *Nat Biotechnol* 22:911–917
  35. Hobman JL (2007) MerR family transcription activators: similar designs, different specificities. *Mol Microbiol* 63:1275–1278