Molecular Plant • Volume 2 • Number 2 • Pages 308–322 • March 2009 RESEARCH ARTICLE

Comparative Genomic Study of the Thioredoxin Family in Photosynthetic Organisms with Emphasis on *Populus trichocarpa*

Kamel Chibani[a,b], Gunnar Wingsle[b], Jean-Pierre Jacquot[a], Eric Gelhaye[a] and **Nicolas Rouhier[a,1]**

a UMR 1136 Nancy University–INRA, Interactions Arbres Microorganismes, IFR 110 GEEF, Faculté des Sciences, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France

b Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, Umeå Plant Science Centre, 90183 Umeå, Sweden

ABSTRACT The recent genome sequencing of *Populus trichocarpa* and *Vitis vinifera*, two models of woody plants, of *Sorghum bicolor*, a model of monocot using C4 metabolism, and of the moss *Physcomitrella patens*, together with the availability of photosynthetic organism genomes allows performance of a comparative genomic study with organisms having different ways of life, reproduction modes, biological traits, and physiologies. Thioredoxins (Trxs) are small ubiquitous proteins involved in the reduction of disulfide bridges in a variety of target enzymes present in all sub-cellular compartments and involved in many biochemical reactions. The genes coding for these enzymes have been identified in these newly sequenced genomes and annotated. The gene content, organization and distribution were compared to other photosynthetic organisms, leading to a refined classification. This analysis revealed that higher plants and bryophytes have a more complex family compared to algae and cyanobacteria and to non-photosynthetic organisms, since poplar exhibits 49 genes coding for typical and atypical thioredoxins and thioredoxin reductases, namely one-third more than monocots such as *Oryza sativa* and *S. bicolor*. The higher number of Trxs in poplar is partially explained by gene duplication in the Trx m, h, and nucleoredoxin classes. Particular attention was paid to poplar genes with emphasis on Trx-like classes called Clot, thioredoxin-like, thioredoxins of the lilium type and nucleoredoxins, which were not described in depth in previous genomic studies.

Key words: Genomic; oxidoreductase; poplar; redox regulation; thioredoxins.

INTRODUCTION

Dithiol–disulfide exchange reactions participate to the post-translational modification of proteins, and they also influence the redox-state of catalytic and structural thiol groups under changing redox homeostasis. Thioredoxins (Trxs), Trx-like proteins, glutaredoxins (Grx), glutathione, as well as cyclophilins, are involved in these reducing reactions. Thioredoxins are ubiquitous small proteins with a redox-active disulfide bridge, defined by the presence of two cysteinyl residues in the 'classical' WC[G/P]PC catalytic site (Eklund et al., 1991). This active site, conserved in most organisms, is considered as a specific Trx signature together with a conserved structural motif called the 'thioredoxin fold', which consists of a central pleated β-sheet constituted by 5 β-strands surrounded by α-helices. More generally, the CxxC/S motif used by thioredoxins and glutaredoxins is essential for the reduction of inter- and intramolecular disulfide bonds and other forms of oxidized cysteines (Fomenko and Gladyshev, 2003; Gelhaye et al., 2004b; Rouhier et al., 2004). It is also present in a wide variety of proteins and enzymes that are either redox-regulated or that bind metals such as zinc-finger transcription factors (Takatsuji, 1998), methionine sulfoxide reductases (Rouhier et al., 2006a), iron sulphur-containing proteins, or cytochromes belonging to electron transfer chains (Rouhier et al., 2007). Changing one of the amino acids of the active site by site-directed mutagenesis disturbs the functional properties of these thiol-disulfide oxidoreductases most likely by modifying their redox potential and the pK_a of the cysteines (Chivers et al., 1997; Mössner et al., 2000). The thioredoxin superfamily is a large multigene family that comprises five major groups of

[1] To whom correspondence should be addressed. E-mail nicolas.rouhier@scbiol.uhp-nancy.fr, fax +33 3 83 68 42 92, tel. +33 3 83 68 42 25.

doi: 10.1093/mp/ssn076
Received 5 September 2008; accepted 21 October 2008

proteins: thioredoxins, glutaredoxins, protein disulfide isomerases (PDI), glutathione peroxidases, and glutathione-S-transferases, all sharing a common structural arrangement.

Plants have an extended Trx system consisting of several Trx isoforms localized in different sub-cellular compartments such as the chloroplast, mitochondria, cytosol and even the nucleus (Gelhaye et al., 2005). On the other hand, the presence of multiple thioredoxins in the same compartment raises the question of their specificity (Jacquot et al., 1997; Meyer et al., 1999). In plants, unlike bacteria and animals, genes encoding thioredoxins have been grouped into six major classes (f, m, x, y, o, and h) (Gelhaye et al., 2004b; Lemaire et al., 2004; Meyer et al., 2007). The f, h, and o types are evolutionarily closer to eukaryotic sequences, whereas m, x, and y types are more closely related to prokaryotic sequences (Meyer et al., 1999; Gelhaye et al., 2005). The thioredoxins f, m, x, and y are localized in the chloroplast, whereas the thioredoxins o are found in mitochondria (Laloi et al., 2001). The thioredoxins h constitute a large disparate group that includes some cytosolic thioredoxin isoforms but also some Trxs containing N-terminal extensions (Gelhaye et al., 2005). As a result, some members of this group are targeted to mitochondria (poplar Trx h2) or are secreted (Trx h5 *Nicotiana* ortholog) (Gelhaye et al., 2004a; Juárez-Díaz et al., 2006). In another case, the N-terminal extension contains a conserved cysteine that is involved in an atypical regeneration mechanism involving glutathione and glutaredoxin but not the 'classical' NADPH thioredoxin reductase (NTR) (Gelhaye et al., 2003; Koh et al., 2008). The function of thioredoxins h is currently being investigated in several plant systems. In cereals, it has been found to promote the mobilization of carbon and nitrogen of the endosperm early in grain germination (Wong et al., 2002; Shahpiri et al., 2008). A Trx h is required during nodule development to reduce the ROS level in soybean roots (Lee et al., 2005). AtTRXh5 is induced in response to oxidative stress conditions and especially during pathogen infection (Laloi et al., 2004; Sweat and Wolpert, 2007).

This classification has then been extended with the characterization of thioredoxin-like proteins that present either a traditional WCGPC active site or varying CxxC/S active sites and disulfide reductase activity. Within this group, the best studied plant proteins are the atypical CDSP32 (chloroplast drought-induced protein of 32 kDa), which are bimodular Trxs with a HCGPC active site (Rey et al., 1998), the NTRC, which contains a NTR module in the N-terminal part linked to a C-terminal Trx module (Pérez-Ruiz et al., 2006) and the tetratricopeptide domain-containing thioredoxins (TDX). The TDX, which possesses an N-terminal tetratricopeptide repeat domain associated with a C-terminal WCGPC motif, interacts with Ssb2, a yeast heat-shock protein 70 chaperone (Vignols et al., 2003). A recent study on the legume model *Medicago truncatula* has also established the existence of two thioredoxin-like proteins called Trxs, which are targeted to the endoplasmic reticulum and have a function in symbiotic interaction in legumes (Alkhalfioui et al., 2008).

Depending on their sub-cellular localization, thioredoxins are reduced by different electron donor systems. In plastids, most Trxs (m, f, x, and y) are reduced via the ferredoxin/thioredoxin reductase (FTR) system with electrons provided by the photosynthetic electron transport chain (Dai et al., 2004; Balmer et al., 2006; Lemaire et al., 2007; Schürmann and Buchanan, 2008). Alternatively, the NTRc fusion protein can provide electrons from NADPH (Pérez-Ruiz et al., 2006). Nevertheless, this protein is apparently not an efficient reductant for other chloroplastic Trxs (Issakidis Bourguet, personal communication). Outside the chloroplast, the Trxs o and h are generally reduced by electrons coming from NADPH via NADPH thioredoxin reductase (NTR), but some of them are reduced by the glutathione/glutaredoxin system (Laloi et al., 2001; Gelhaye et al., 2003; Reichheld et al., 2005, 2007). Several recent reviews have compiled data about the functions of thioredoxins in plants and this will not be described here (Vieira dos Santos and Rey, 2006; Lemaire et al., 2007; Rouhier et al., 2008; Schürmann and Buchanan, 2008).

Comparative genomic analyses highlighted the presence of expanded gene families encoding Grx and Trxs in higher plants, algae, and mosses (Lemaire and Miginiac-Maslow, 2004; Rouhier et al., 2004, 2006b; Gelhaye et al., 2004b, 2005; Meyer et al., 2005, 2006, 2007; Florencio et al., 2006; Nuruzzaman et al., 2008), whereas non-photosynthetic organisms contain only a limited number of those genes and proteins. To date, these analyses indicate that there are about 40 potential Trx genes in *Arabidopsis thaliana*, 30 in *Oryza sativa*, 11 in *Chlamydomonas reinhardtii*, and 22 in *Populus trichocarpa*. Nevertheless, some thioredoxin or thioredoxin-like proteins have sometimes been excluded due in part to false automatic annotation or also by using a too strict definition of thioredoxin (i.e. consider only the WCGPC sequences and not all variants of the active site) and this has made some analyses incomplete. Thus, a clear and global classification of the Trx family is still not fully accomplished, especially because an increasing number of plant and algal genomes is now available and also because new experimental data concerning these disulfide reductases are revealing a growing complexity. The present Trx classification is essentially based on the gene structure, on the protein primary structure and especially on the active site sequence and on the sub-cellular localization.

This paper presents a large and hopefully complete overview of the Trx gene family including the conventional Trx h, m, f, x, y, o, atypical thioredoxins such as Trx-like, Lilium-type, Clot, CDSP32, TDX, HCF164, Cf-9 interacting thioredoxin (CiTrx), and nucleoredoxin (Nrx) and their reductases (NTR, FTR beta chain) (Lennartz et al., 2001; Rivas et al., 2003). After manual gene curation and annotation, this in-silico analysis performed primarily with *P. trichocarpa* served as a basis for a comparative genomic study of thioredoxins and thioredoxin-like proteins using several sequenced photosynthetic organisms possessing different organization and physiology, namely higher plants (*A. thaliana*, *O. sativa*, *V. vinifera*, *S. bicolor*),