

OPEN

Analysis of *Arabidopsis thaliana* Redox Gene Network Indicates Evolutionary Expansion of Class III Peroxidase in Plants

Raffael Azevedo de Carvalho Oliveira^{1,2}, Abraão Silveira de Andrade¹, Danilo Oliveira Imparato¹, Juliana Gabriela Silva de Lima², Ricardo Víctor Machado de Almeida², João Paulo Matos Santos Lima^{1,2}, Matheus Augusto de Bittencourt Pasquali^{3,4,5} & Rodrigo Juliani Siqueira Dalmolin^{1,2*}

Reactive oxygen species (ROS) are byproducts of aerobic metabolism and may cause oxidative damage to biomolecules. Plants have a complex redox system, involving enzymatic and non-enzymatic compounds. The evolutionary origin of enzymatic antioxidant defense in plants is yet unclear. Here, we describe the redox gene network for *A. thaliana* and investigate the evolutionary origin of this network. We gathered from public repositories 246 *A. thaliana* genes directly involved with ROS metabolism and proposed an *A. thaliana* redox gene network. Using orthology information of 238 Eukaryotes from STRINGdb, we inferred the evolutionary root of each gene to reconstruct the evolutionary history of *A. thaliana* antioxidant gene network. We found two interconnected clusters: one formed by SOD-related, Thiol-redox, peroxidases, and other oxido-reductase; and the other formed entirely by class III peroxidases. Each cluster emerged in different periods of evolution: the cluster formed by SOD-related, Thiol-redox, peroxidases, and other oxido-reductase emerged before opisthokonta-plant divergence; the cluster composed by class III peroxidases emerged after opisthokonta-plant divergence and therefore contained the most recent network components. According to our results, class III peroxidases are in expansion throughout plant evolution, with new orthologs emerging in each evaluated plant clade divergence.

Reactive oxygen species (ROS) are byproducts of normal metabolism of aerobic organisms, being produced in a constitutive and not controlled manner during photosynthesis in plants and during respiration in most living beings. External agents such as UV radiation, pathogens, extreme temperatures, pollution, and xenobiotics can contribute to ROS production. Organisms have developed antioxidant systems during evolution to avoid ROS toxicity. Plants are constantly exposed to stressing agents, mainly due to their sessile nature. Many of those agents can imbalance ROS production, leading to oxidative stress¹. In spite of their toxic trait, ROS can be useful in plants acting as signaling molecules^{2–4} and in defense against pathogens^{5,6}. Enzymes such as RBOH (Respiratory Burst Oxidase Homologue), a homologous enzyme of mammalian NADPH-oxidase, as well as some peroxidases, are responsible for producing H₂O₂ in plants^{7,8}. H₂O₂ has several roles in plant physiology as signaling for cell expansion⁹, seed germination¹⁰, stomatal closure¹¹, and hormone production¹². Additionally, H₂O₂ has an important role in cell wall reinforcement since it is involved with lignin synthesis⁷. The balance between ROS production and scavenging should be spatially and temporally regulated in plants to avoid oxidative damage and to allow ROS-mediated cell signaling^{13,14}.

Plants have developed complex antioxidant machinery comprising enzymatic and non-enzymatic mechanisms to protect against ROS overproduction¹⁵. The non-enzymatic antioxidant machinery involves phenolic compounds,

¹Bioinformatics Multidisciplinary Environment - IMD, Federal University of Rio Grande do Norte, Natal, Brazil.

²Department of Biochemistry, Federal University of Rio Grande do Norte, Natal, Brazil. ³Institute of Tropical Medicine, Federal University of Rio Grande do Norte, Natal, Brazil. ⁴Food Engineering Unit, UAEAL, UFCG, Campina Grande, Brazil. ⁵Graduate Program in Natural Resources, PPGRN, UFCG, Campina Grande, Brazil. *email: rodrigo.dalmolin@imd.ufrn.br

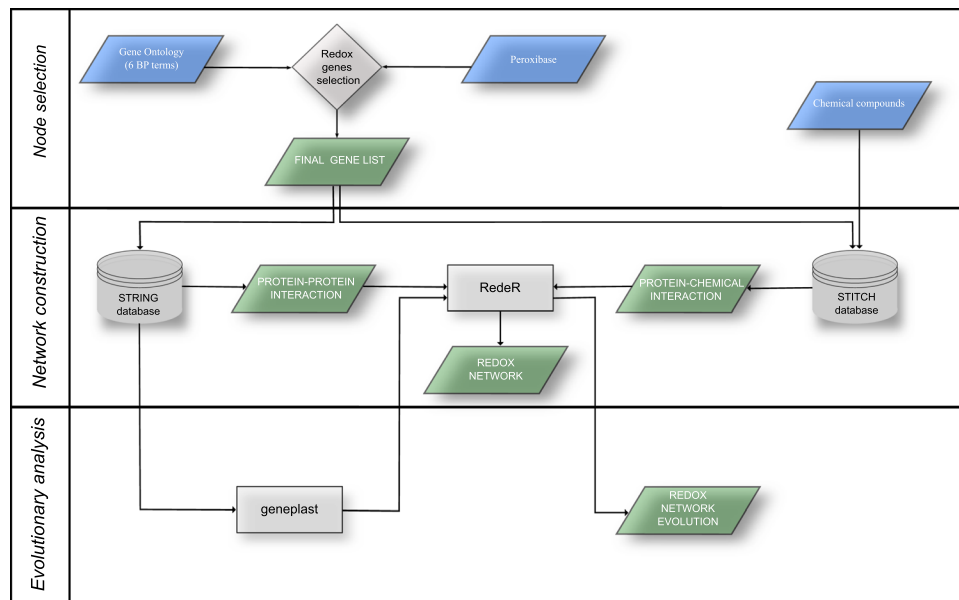


Figure 1. Analysis workflow. Redox genes were selected from Gene Ontology and PeroxiBase and then manually curated (Node selection). Protein-protein interactions were obtained from STRING database, protein-chemical interactions were obtained from STITCH database, and the final network was handled using RedeR Bioconductor package (Network construction). The evolutionary root of each *A. thaliana* antioxidant gene was inferred using geneplast Bioconductor package (Evolutionary analysis).

ascorbate, tocopherols, carotenoids, flavonoids, abscisic and jasmonic acids, forming a defense line against ROS¹⁶. The non-enzymatic defenses act coordinately with several antioxidant enzymes, such as catalase (CAT), peroxidases (PER), superoxide dismutases (SOD), and other proteins with no catalytic activity, to maintain redox homeostasis¹⁷. Both enzymatic and non-enzymatic defenses work together in many known reactions, like the glutathione-ascorbate cycle, characterizing a redox-buffer system¹⁸. Peroxidases (EC 1.11.1.X) have a central role in plant metabolism. Besides hydrogen peroxide scavenging, peroxidases can act in H₂O₂ generation and are essential players in auxin catabolism, suberization, programmed cell death, defenses against pathogens, and lignification of cell wall^{19–21}, which might have been a critical evolutionary step on land colonization. Antioxidant and ROS-generating enzymes work together in a redox network to coordinately maintain redox homeostasis in plants^{22,23}.

The evolutionary origin of plant redox systems is a matter of debate. Antioxidant defenses are thought to date back Paleoproterozoic Great Oxidation Event or earlier^{24,25}. Many antioxidant enzymes found in plants have orthologs in fungi and metazoan^{25–27}. Plants, however, have particularities regarding redox metabolism since ROS perform many specific functions in plants as mentioned before. It is reasonable to think that plant redox gene network differs from other eukaryotic groups and those differences might be crucial, for instance, to plant acclimation to stressing conditions²⁷. There are three peroxidase groups in eukaryotes: (i) class I peroxidases, widely distributed within eukaryotes; (ii) class II peroxidases, exclusive to fungi; and (iii) class III peroxidases, exclusive to plants^{21,26}. Several class III peroxidase isozymes have been described in different angiosperms such as *A. thaliana*, *Oryza sativa*, and *Zea mays*²¹. The vast number of plant-exclusive peroxidases found in angiosperms brings clues about its importance in the expansion of redox system in plants.

Here, we propose the plant redox gene network using the model organism *Arabidopsis thaliana*. Using PeroxiBase, Gene Ontology, and TAIR database, a reference database for *A. thaliana*, we identified 323 genes directly involved with ROS metabolism, being 246 of them present in the network main connected component. The network, based on protein-protein and protein-chemical interactions, presented two clusters: one composed by Thiol-redox proteins, SOD-related proteins, peroxidases, and other oxido-reductase, and the other formed exclusively by class III peroxidases. We used orthology information regarding 238 eukaryotes to infer the evolutionary root of each gene to reconstruct the evolutionary history of the *A. thaliana* redox gene network. We found class III peroxidases as the most recent components of the network, showing that class III peroxidases cluster is in expansion in plants.

Results

A. thaliana redox gene network. *A. thaliana* redox genes were selected according to two criteria: using GO consortium and PeroxiBase (Fig. 1). For the first criteria, we selected genes from six GO terms associated with redox metabolism, and for the second criteria, we selected *A. thaliana* genes annotated in PeroxiBase. The merged gene list included 323 genes. Among them, 246 were connected to the network main connected component, and 77 genes presented no connection when following the parameters described above, including *PER48*, *PER1*, and *PER62* (Supplementary Table S3). The final network includes 246 genes and 15 chemical compounds directly involved with the redox enzymes, being their substrates or products (Fig. 2a).

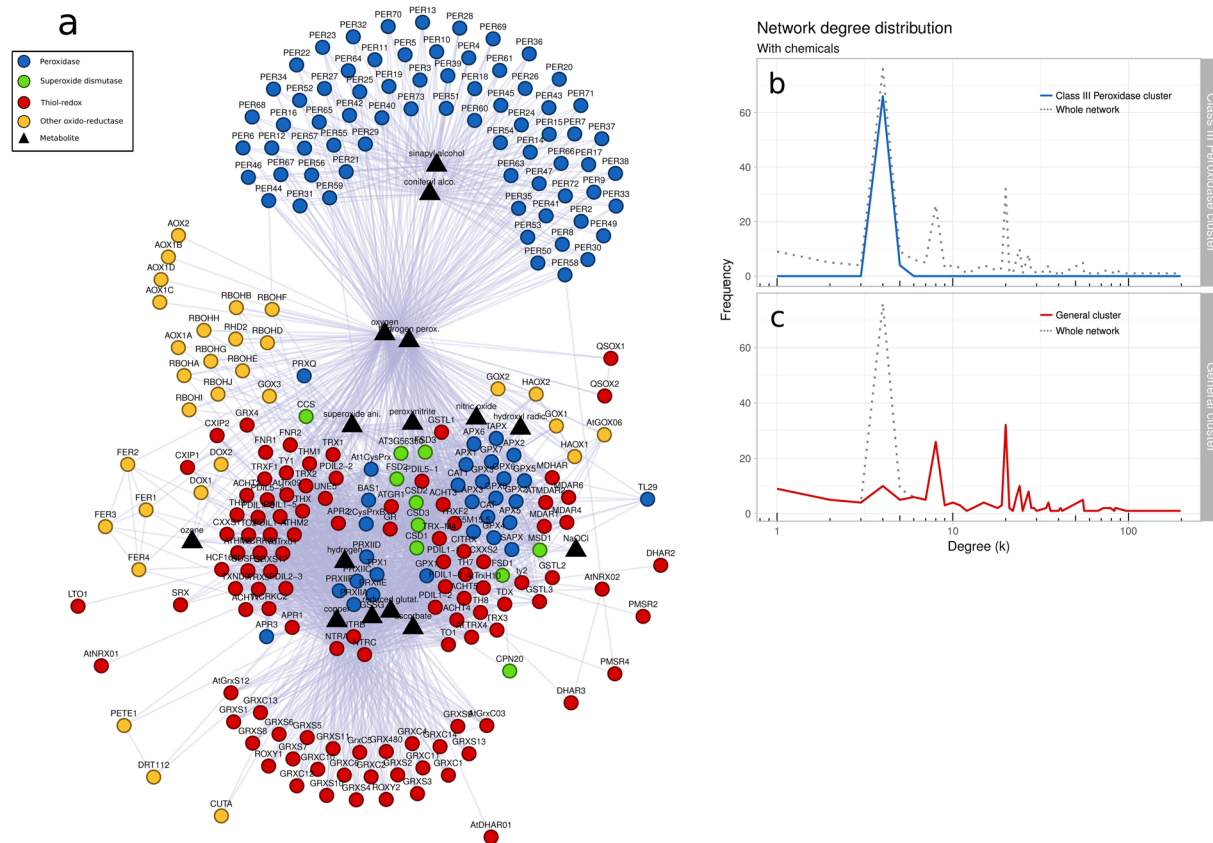


Figure 2. *Arabidopsis thaliana* redox gene network. (a) Protein-protein interaction and protein-chemical interaction among 246 *A. thaliana* redox genes (circle nodes) and 15 antioxidant enzyme substrates and/or products (triangle nodes). Node color indicates protein class: peroxidase (blue), superoxide dismutase (green), thiol-redox (red), and other oxido-reductase (yellow). Chemical nodes are shown in black. (b) Degree distribution (red line) of proteins from *class III peroxidase cluster*. The dotted line shows degree distribution of the whole network, including chemical compounds. (c) Degree distribution (blue line) of proteins from the *general cluster*. The dotted line shows degree distribution of the whole network, including chemical compounds.

The result is a metabolic network involving protein-protein interaction and protein-chemical interaction. Antioxidant systems are described to work coordinately to maintain physiological redox status, acting as an antioxidant-defense network^{22,28,29}. Those properties can be clearly observed in Fig. 2a by the entanglement of the network components.

Genes were classified according to their function: 106 thiol-redox, 100 peroxidases, 10 SOD-related genes, and 30 other oxido-reductase (including RBOH isoforms, ferritin isoforms, and alternative oxidases). According to Fig. 2a, *A. thaliana* redox gene network consists of two clusters with distinct functions. The cluster on the bottom is formed by SOD-related, thiol-redox, peroxidases, and other oxido-reductase, totaling 176 genes. From now on we will refer to the bottom cluster as *general cluster*. The cluster on the top is composed entirely by class III peroxidase genes, totaling 70 genes. We will refer to the top cluster as *class III peroxidase cluster*. The nodes in the *general cluster* are very interconnected while genes in the *class III peroxidase cluster* are linked mainly by the substrates/products O_2 , H_2O_2 , coniferyl, and sinapyl (Fig. 2b,c and Supplementary Fig. S1). O_2 and H_2O_2 work as hubs linking the two clusters. Hydrogen peroxide is known to perform several physiological functions in plants and its central position in the network denotes its importance in redox plant metabolism.

Evolutionary Origin Of *A. thaliana* Redox Gene Network. While proteins within the *general cluster* are found in other eukaryotes (e.g., SOD, GPX, and other peroxidases), class III peroxidases are exclusive to plants. We then assessed the evolutionary root of each redox gene to investigate the evolutionary origin of *A. thaliana* redox gene network. The idea is to identify the most ancient genetic archetype of each *A. thaliana* redox gene that has been vertically inherited. To do that, we looked for *A. thaliana* redox gene orthologs in 238 eukaryotes (see Supplementary Fig. S2). According to our results, the majority of the genes within the *general cluster* emerged before Viridiplantae divergence (Fig. 3). Except for two oxido-reductase (*DOX1* and *DOX2*) and two thiol-redox (*DHAR2* and *DHAR01*), all the other thiol-redox protein-coding genes, all glutathione peroxidases, ascorbate peroxidases, catalases, thioredoxin-dependent peroxidases, superoxide dismutases isoforms (Mn, Cu-Zn, and Fe), and their cofactor-supplier proteins (*ATCCS* and *CPN20*) compose the ancient portion of the network. *Naegleria gruberi* and *Giardia lamblia* are the most phylogenetically distant species from *A. thaliana* in our species set (Supplementary Fig. S2). At total, 146 proteins of the *general cluster* have orthologs rooted at the

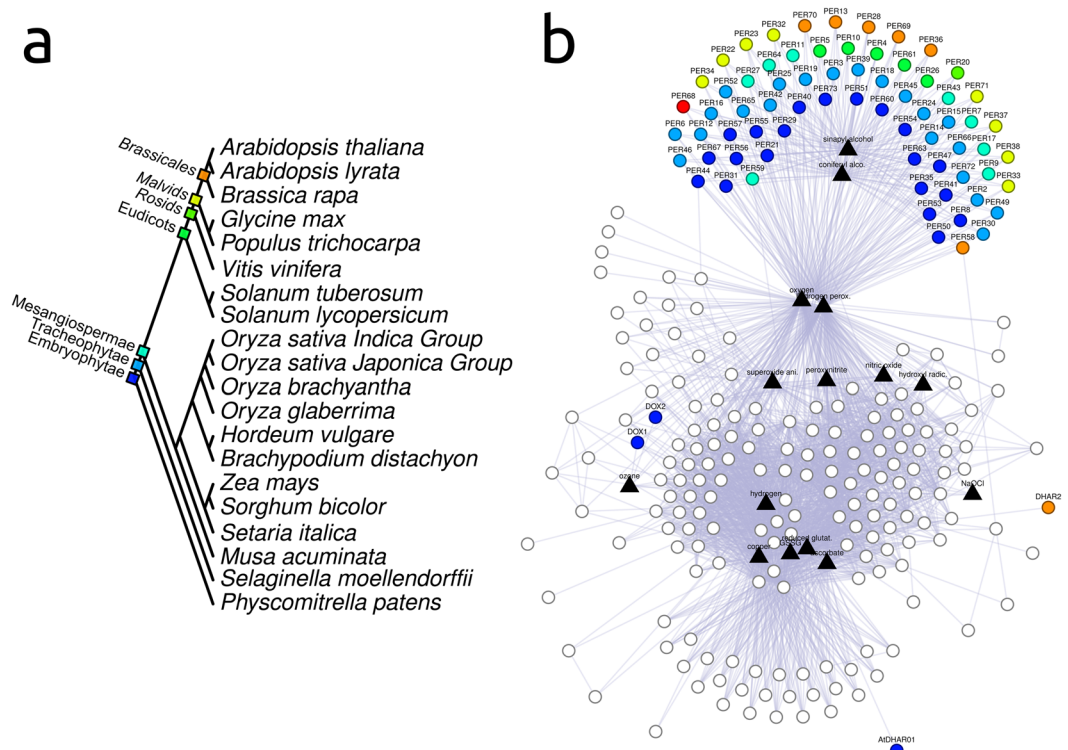


Figure 3. Evolutionary root of *A. thaliana* class III peroxidases orthologs. **(a)** Evolutionary tree of 20 land plants used in the rooting analysis. Colored diamonds indicate the last common ancestral of each taxonomic group. **(b)** *A. thaliana* redox gene network showing the inferred evolutionary root of each gene on the network. Node color represents the evolutionary root of each gene, according to the evolutionary tree shown in A. Genes which the inferred evolutionary root was placed before plant divergence are shown in white. It was not possible to determine the evolutionary root of *PER68* (shown in red).

divergence among *Naegleria gruberi*/*Giardia lamblia* and *A. thaliana*. It represents that more than 80% of the *general cluster* is rooted at the origin of eukaryotes. Additionally, only 5% of the *general cluster* arrived after the divergence among opisthokonta and plants. These results suggest that *general cluster* is the ancient part of *A. thaliana* redox network (for a detailed analysis of the *general cluster* evolution, please see Supplementary Figs S2 and S3).

On the other hand, all *A. thaliana* class III peroxidases orthologs emerged at Viridiplantae divergence or afterwards. It is important to mention that *PER68* is not assigned in any orthologous group (OG) on STRINGdb, so it was not possible to determine its evolutionary root following the strategy used here. This might be a result of the COG algorithm, which needs a triangulation among at least three species to build a cluster of orthologs. Accordingly, proteins found in less than three species do not take part in an COG. We found 20 class III peroxidase genes rooted at Viridiplantae divergence, 21 rooted at tracheophytes divergence, 8 at angiosperms divergence, 5 at eudicots divergence, 1 at rosids divergence, 8 at malvids divergence, and 6 at brassicales divergence (Fig. 3 and Supplementary Fig. S4).

Class III Peroxidase Abundance. As mentioned before, class III peroxidases are very abundant in *A. thaliana*. We evaluate the abundance of class III peroxidases orthologs in the different plant taxonomic groups available in our dataset. We assessed two sets of orthologous groups, *A. thaliana* class III peroxidases OGs and *Oryza sativa* class III peroxidases OGs (see methods), to cover the OGs distribution in the two main monophyletic groups of angiosperms: monocots and eudicots. Figure 4 shows the average number of class III peroxidases orthologs per species in eudicots (8 species), monocots (10 species), *Selaginella moellendorffii*, and *Physcomitrella patens*. *Physcomitrella patens*, a non-vascular plant, have roughly half of class III peroxidases when compared with vascular plants, including the non-seed vascular plant *Selaginella moellendorffii* (Fig. 4). To avoid a possible bias due to the differences in the number of genes in each evaluated species, we normalized class III peroxidases OGs abundance by the number of genes of each species, finding basically the same results: *P. patens* have roughly half of class III peroxidases when compared with vascular plants even taking into account the number of genes in each species (Supplementary Fig. S4). The distribution pattern is similar in OGs shared by *A. thaliana* and *O. sativa*. Angiosperms have more class III peroxidases per species when compared to *S. moellendorffii*. However, it is difficult to draw conclusions about class III peroxidases distribution since we have only one Moss and one Lycopodiopsida (*i.e.*, *P. patens* and *S. moellendorffii*, respectively) in our dataset. Focusing on angiosperms, we can observe a high number of class III peroxidases orthologs found in eudicots which are not found in monocots and vice versa. Additionally, we can observe a higher number of class III peroxidases in monocots when compared to eudicots (Fig. 4).

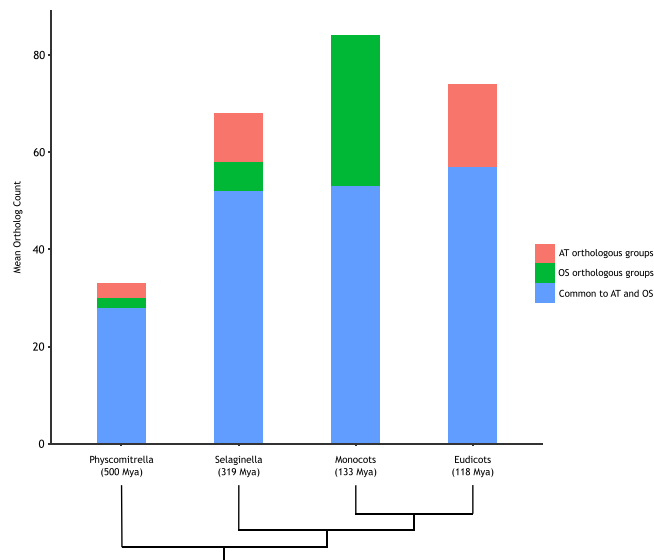


Figure 4. Class III peroxidase orthologs abundance. Bars indicate the number of class III peroxidase isozymes by species of each taxonomic group (Bryophyta - *Physcomitrella patens*, Lycopodiophyta - *Selaginella moellendorffii*, Eudicots - 8 species, and Monocots - 10 species). Colors indicate whether the OG have *A. thaliana* orthologs (AT orthologous groups - pink), *O. sativa* orthologs (OS orthologous groups - green) or orthologs from both species (Common to AT and OS - blue). The dendrogram at the bottom shows the evolutionary relationship among the groups. Estimated divergence times are shown.

Using PeroxiBase information, we redid the analysis including *Marchantia polymorpha*, a species not represented in the dataset used here. According to PeroxiBase, *M. polymorpha* has 167 class III peroxidase distributed in 23 OGs, 8 of which containing orthologs from both *A. thaliana* and *O. sativa*, 1 OG containing orthologs from *A. thaliana* (with no *O. sativa* ortholog) and 1 OG containing orthologs from *O. sativa* (with no ortholog of *A. thaliana*). We used *M. polymorpha* OGs inferred by PeroxiBase since *M. polymorpha* is not among the 238 eukaryotes present in STRING. At total, *M. polymorpha* has 23 class III peroxidases that were identified as ortholog of *A. thaliana* class III peroxidases in PeroxiBase. When compared to *Physcomitrella patens*, *M. polymorpha* has a similar number of class III peroxidases which have orthologs in *A. thaliana* and/or *O. sativa*. However, there are too few species of basal plants in our sample to allow more conclusive results (Supplementary Fig. S5).

Discussion

Plants are constantly exposed to stressing conditions and have developed mechanisms to deal with stress compounds. Oxidant compounds are recognized as being one of the most important stressing agents in plants^{30,31}. On the other hand, ROS are also extensively known to take part in several physiological processes, being indispensable for plants^{13,23,32}. Plant redox system should be able to finely tune cellular redox status to deal with the double-edged sword behavior of oxidant compounds correctly²³. Here, we reconstructed the redox network of *A. thaliana* and tracked down the evolutionary roots of its genes to identify the origin of genetic redox system in plants.

In order to depict the *A. thaliana* redox network, we gathered canonical protein-protein interaction (PPI) and protein-chemical interaction involving the main redox enzyme substrates and/or products. A similar strategy has been used by Gelain and collaborators to describe the human antioxidant gene network²⁸. According to the authors, antioxidant proteins act coordinately, forming an antioxidant buffer. Therefore, it is necessary to evaluate their connections among each other, as well as the connections with their respective substrates and/or products, to get a complete perspective of the redox gene network²⁸. In 2004, Mittler and collaborators proposed a reactive oxygen gene network of plants²². They gathered both ROS-producer and ROS-scavenger protein-coding genes that are part of an extensive redox system. They proposed that the plant ROS network would have about 150 genes, including all ROS-producing RBOH isozymes and other oxidases. Here, we found 60% more genes when compared with Mittler's results. It is important to note that the work conducted by Mittler and collaborators did not include class III peroxidases, in spite of the authors have assumed that the so-called *classical plant peroxidases* act in hydrogen peroxide scavenging²². Although some authors have described the genes involved with redox homeostasis in plants, the network-based analysis proposed here allowed the identification of *class III peroxidase cluster*, which is specific to plants and consequently evolved after opisthokonta-plant divergence. According to our results, modern plants share the redox enzymatic core (*e.g.*, SOD, glutathione peroxidases, and other peroxidases and oxido-reductases) with other Eukaryotes. This agrees with literature where antioxidant enzymes, such as SOD, CAT, and several Thiol-redox, have been described as ancient^{21,33,34}. Some of those genes, like the Fe-SOD, could be found in Archea³⁵ and other proteins known as antioxidant in modern plants emerged before the great oxidation event^{25,36}. Even ascorbate peroxidase (APX) enzymes, which are central to plant redox homeostasis, are known to have evolved before plant divergence and are well described in ancestral taxonomic groups such as Trypanosomatidae²⁷.

On the other hand, we showed that *A. thaliana* class III peroxidase orthologs have emerged at opisthokonta-plant divergence and the group has increased during plant evolutionary expansion since it was possible to identify the emergence of new class III peroxidases COGs in each evaluated plant clade divergence. Inupakutika and collaborators investigated the evolution of plant ROS gene network, focusing SOD, APX, and RBOH genes. They evaluated 11 genes, involving ROS generator and ROS scavenger genes, in 33 species, including 30 eukaryotes and 6 plants. The authors found SOD, peroxiredoxins (PRXR), CAT, and APX as ancient genes, shared by many taxonomic groups²⁵. We have very similar results regarding the antioxidant core (*i.e.*, the *general cluster*). However, we used a different methodology designed to evaluate entire genetic systems, allowing to investigate the evolutionary origin of all 246 *A. thaliana* redox genes present in the network main connected component and using 238 eukaryotes as sample space.

Besides the importance of recognizing the core antioxidant enzymes (*i.e.*, SOD, CAT, PRXRs, APX, other class I peroxidases, and other oxido-reductases) as ancient, our results suggest class III peroxidase as particularly important to plants. Not just because they are plant-exclusive, but also due to their evolutionary expansion throughout plant phylogeny (Figs 3 and 4). Studies made in Chinese pear and maize shows that class III peroxidase genes underwent positive selection^{37,38}. Additionally, the fact that 72 isozymes fit in 44 different orthologous groups also corroborates with the complexity related to this class of protein³⁹. Our result corroborates with many evolutionary analyses made for class III peroxidases^{40–42}, as well as with the theory that some class III peroxidases isozymes emerged more likely at the beginning of land colonization by plants⁴³, in spite of some works having described class III peroxidases members in algae^{26,44}.

Class III peroxidases are intimately related with H₂O₂ dynamics in plants since this class of peroxidases works in H₂O₂ detoxification and formation, including those which have lignification as primary function⁴⁵. H₂O₂ is crucial to plant redox signaling, defense against pathogens, and is potentially toxic as ROS⁴⁶. Class III peroxidases are intimately involved with lignin polymerization, but we found class III peroxidases orthologs that emerged with Embryophyta divergence. Accordingly, it is difficult to say if class III peroxidases emergence was related to lignification or to deal with increased ROS in an atmosphere with high oxygen concentration during land colonization by plants^{8,44}. Only at Tracheophyta divergence, other class III peroxidase emergence seems to be related to lignification and suberization processes⁸.

The results shown here improve the comprehension of the evolutionary dynamics of the redox gene network in plants and suggest an essential role of class III peroxidases on plant ROS metabolism since *class III peroxidase cluster* seems to evolve continuously throughout plant evolution. It is important to mention that enzymatic antioxidant defenses are just part of the whole redox machinery. Plants have complex biochemical pathways to produce non-enzymatic antioxidant compounds, such as carotenoids and flavonoids, and those molecules are crucial to redox metabolism. The genetic network proposed here allowed to identify classical antioxidant enzymes, such as SOD, CAT, glutathione peroxidases (GPX), etc., as ancient and class III peroxidases as the novelty regarding redox plant metabolism. Taking together, the results herein suggest that class III peroxidases have a pivotal role in plant redox homeostasis and the expansion of redox network throughout plant evolution are coincident with the continuous emergence of class III peroxidases new orthologs.

Methods

Arabidopsis redox gene selection. *A. thaliana* redox genes were selected using GO consortium⁴⁷ and PeroxiBase⁴⁸, only including genes also annotated on TAIR database (<https://www.arabidopsis.org/>)⁴⁹. We selected *A. thaliana* redox genes from six GO terms associated with redox metabolism: (i) *oxygen species metabolic process* (GO:0072593), (ii) *cell redox homeostasis* (GO:0045454), (iii) *cellular response to reactive oxygen species* (GO:0034614), (iv) *response to inorganic substance* (GO:0010035), (v) *oxidation-reduction process* (GO:0055114), and (vi) *response to oxidative stress* (GO:0006979). Gene function was checked in TAIR database (<https://www.arabidopsis.org/>)⁴⁹. Genes were then manually curated according to the criteria previously established by Gelain and collaborators²⁸. Briefly, the selection includes genes whose products are described as primarily antioxidant scavengers as well as genes indispensable to antioxidant enzymes activity. The objective was to include only gene products that directly deal with oxidants. The resulting gene set was manually curated to identify 208 genes directly involved with redox homeostasis (Supplementary Table S1 and S3).

We also collected *A. thaliana* genes annotated in PeroxiBase (now named as RedoxiBase, since the name has changed during the editorial processes). PeroxiBase is a manually curated database dedicated mainly to peroxidase superfamilies, including other proteins known to deal with ROS⁴⁸. Gene function was checked in TAIR database and the final PeroxiBase gene list content reached 293 genes (Supplementary Table S1 and S3).

The final gene set includes 323 genes (see Supplementary Table S1 and S3). The genes were classified into four categories according to the TAIR database description: thiol-redox, peroxidase, SOD, and other oxido-reductase.

Redox gene network. *A. thaliana* redox gene network was built using STRING database version 10.5 (<https://version-10-5.string-db.org/>)⁵⁰ with input options ‘databases’, ‘experiments’, and 0.400 confidence level. STRING integrates different curated databases containing information on direct and indirect functional protein–protein associations. The search was carried out using TAIR id obtained from TAIR database and the connections among the 323 redox genes were collected. In addition to redox genes, the network also included chemical compounds that are substrates or products of the antioxidant enzymes. In total, were included 15 chemical compounds (see Supplementary Table S2). The connections involving the 15 chemical compounds and the 323 redox genes were retrieved from STITCH database version 5.0 (<http://stitch.embl.de/>)⁵¹ with input options ‘databases’, ‘experiments’, and 0.400 confidence level. STITCH integrates information regarding interactions between proteins and small molecules. 77 genes presented no connection when following the parameters described above (Supplementary Table S3). The resultant network includes protein–protein interactions as well as protein–chemical interactions among 246 *A. thaliana* redox genes and the 15 redox-related chemical compounds

(Supplementary Fig. S2). The final network was handled using RedeR, an R package for interactive visualization and manipulation of networks⁵².

Evolutionary rooting analysis. The evolutionary rooting analysis was carried out using geneplast, an R package designed to perform evolutionary rooting based on the distribution of orthologous groups, set with default parameters⁵³. The method uses orthology information to infer the probability that a feature had been present in each Last Common Ancestor of a given species tree. The problem was previously addressed by Mirkin⁵⁴ and Castro⁵⁵. Concerning our investigation, the task is: to find its earliest ortholog in the eukaryote phylogeny for each orthologous group associated with the *A. thaliana* redox network. Orthology information of the 246 *A. thaliana* redox genes was retrieved from clusters of orthologous groups (COGs) from STRINGdb v10.5⁵⁰. The analysis was performed using 238 eukaryotes available in STRINGdb (Supplementary Fig. S1). In total, 119 orthologous groups were identified for 245 of the 246 genes present in the network main connected component (Supplementary Table S1). The evolutionary trees describing the phyletic pattern of each redox gene are presented in Supplementary Material (see Supplementary Figs S6 to S125).

Class III peroxidase abundance. The class III peroxidase genes abundance was calculated taking into account the orthologous genes from *A. thaliana* and/or *O. sativa* class III peroxidase genes in the 20 land plant species present on STRING database. *A. thaliana* class III peroxidase orthologous genes were collected as described above. From the original 119 redox *A. thaliana* OGs, 44 were annotated as having class III peroxidase proteins (see Supplementary Table S4). To identify *O. sativa* class III peroxidase OGs not shared with *A. thaliana*, we searched for *O. sativa* class III peroxidase genes in PeroxiBase⁴⁸ and then translated to STRING format using UniProt⁵⁶ and Ensembl Biomart⁵⁷. At total, 53 *O. sativa* class III peroxidase OGs were used to retrieve the orthologous groups from STRING database (see Supplementary Table S5), which 29 were shared with *A. thaliana*.

Data availability

The datasets analyzed during the current study are publicly available as Supplementary Material. Any other data/protocol is available from the corresponding author on request.

Received: 25 January 2018; Accepted: 16 August 2019;

Published online: 31 October 2019

References

- Kilian, J. *et al.* The AtGenExpress global stress expression data set: Protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J.* **50**, 347–363 (2007).
- Jaspers, P. & Kangasjärvi, J. Reactive oxygen species in abiotic stress signaling. 405–413 (2009).
- Foyer, C. H. & Noctor, G. Redox signaling in plants. *Antioxid. Redox Signal.* **18**, 2087–90 (2013).
- Foyer, C. H. & Noctor, G. Stress-triggered redox signalling: What's in pROSpect? *Plant, Cell Environ.* 1–14 <https://doi.org/10.1111/pce.12621> (2015).
- Peleg-Grossman, S., Volpin, H. & Levine, A. Root hair curling and Rhizobium infection in *Medicago truncatula* are mediated by phosphatidylinositol-regulated endocytosis and reactive oxygen species. *J. Exp. Bot.* **58**, 1637–1649 (2007).
- Nanda, A. K., Andrio, E., Marino, D., Pauly, N. & Dunand, C. Reactive oxygen species during plant-microorganism early interactions. *J. Integr. Plant Biol.* **52**, 195–204 (2010).
- O'Brien, J. A., Daudi, A., Butt, V. S. & Paul Bolwell, G. Reactive oxygen species and their role in plant defence and cell wall metabolism. *Planta* **236**, 765–779 (2012).
- Weng, J. & Chapple, C. The origin and evolution of lignin biosynthesis. *New Phytol.* **187**, 273–285 (2010).
- Foreman, J. *et al.* Reactive oxygen species produced by NADPH oxidase regulate plant cell growth. *Nature* **422**, 442–446 (2003).
- Barba-Espín, G. *et al.* Understanding the role of H₂O₂ during pea seed germination: A combined proteomic and hormone profiling approach. *Plant, Cell Environ.* **34**, 1907–1919 (2011).
- Miller, G., Suzuki, N., Ciftci-Yilmaz, S. & Mittler, R. Reactive oxygen species homeostasis and signalling during drought and salinity stresses. *Plant, Cell Environ.* **33**, 453–467 (2010).
- Barba-Espín, G. *et al.* Interaction between hydrogen peroxide and plant hormones during germination and the early growth of pea seedlings. *Plant, Cell Environ.* **33**, 981–994 (2010).
- Bailey-Serres, J. & Mittler, R. The roles of reactive oxygen species in plant cells. *Plant Physiol.* **141**, 900191 (2006).
- Dietz, K.-J., Mittler, R. & Noctor, G. Recent progress in understanding the role of reactive oxygen species in plant cell signaling. *Plant Physiol.* **171**, 1535–1539 (2016).
- Pandhair, V. & Sekhon, B. Reactive oxygen species and antioxidants in Plants: An Overview. *J. Plant Biochem. Biotechnol.* **15**, 71–78 (2006).
- Apel, K. & Hirt, H. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu. Rev. Plant Biol.* **55**, 373–399 (2004).
- Praduman, Y., Sunil, K., Prathap, R., Taruna Y, I. M. Oxidative Stress and Antioxidant Defense System in Plants. *Biotechnol. Plant Biotechnol.* **2** (2015).
- Foyer, C. H. & Noctor, G. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol.* **155**, 2–18 (2011).
- Cosio, C. & Dunand, C. Specific functions of individual class III peroxidase genes. *J. Exp. Bot.* **60**, 391–408 (2009).
- Almagro, L. *et al.* Class III peroxidases in plant defence reactions. *J. Exp. Bot.* **60**, 377–390 (2009).
- Obinger, C. & Zámocký, M. *Molecular Phylogeny of Heme Peroxidases. Biocatalysis Based on Heme Peroxidases* https://doi.org/10.1007/978-3-642-12627-7_2 (2010).
- Mittler, R., Vanderauwera, S., Gollery, M. & Van Breusegem, F. Reactive oxygen gene network of plants. *Trends Plant Sci.* **9**, 490–498 (2004).
- Mittler, R. ROS Are Good. *Trends Plant Sci.* **22**, 11–19 (2017).
- Gacesa, R., Dunlap, W. C., Barlow, D. J., Laskowski, R. A. & Long, P. F. Rising levels of atmospheric oxygen and evolution of Nrf2. *Sci. Rep.* **6**, 27740 (2016).
- Inupakutika, M. A., Sengupta, S., Devireddy, A. R., Azad, R. K. & Mittler, R. The evolution of reactive oxygen species metabolism. *J. Exp. Bot.* **67**, 5933–5943 (2016).
- Lazzarotto, F., Turchetto-Zolet, A., Carina & Margis-Pinheiro, M. Revisiting the Non-Animal Peroxidase Superfamily. *Trends Plant Sci.* **20**, 807–813 (2015).

27. Maruta, T., Sawa, Y., Shigeoka, S. & Ishikawa, T. Diversity and Evolution of Ascorbate Peroxidase Functions in Chloroplasts: More Than Just a Classical Antioxidant Enzyme? *Plant Cell Physiol.* **57**, 1377–1386 (2016).
28. Gelain, D. P. *et al.* A systematic review of human antioxidant genes. *Front. Biosci.* **14**, 4457–4463 (2009).
29. Dalmolin, R. J., Gelain, D. P., Klamt, F., Castro, M. A. A. & Moreira, J. C. F. Transcriptomic analysis reveals pH-responsive antioxidant gene networks. *Front. Biosci.* **54**, 1556–1567 (2012).
30. De Gara, L., Locato, V., Dipierro, S. & de Pinto, M. C. Redox homeostasis in plants. The challenge of living with endogenous oxygen production. *Respir. Physiol. Neurobiol.* **173**, S13–S19 (2010).
31. Loiacono, F. V. & De Tullio, M. C. Why We Should Stop Inferring Simple Correlations Between Antioxidants and Plant Stress Resistance: Towards the Antioxidomic Era. *OMICS: A Journal of Integrative Biology* **16**, 160–167 (2012).
32. Mittler, R. *et al.* ROS signaling: The new wave? *Trends Plant Sci.* **16**, 300–309 (2011).
33. Zamocky, M., Furtmüller, P. G. & Obinger, C. Evolution of catalases from bacteria to humans. *Antioxid. Redox Signal.* **10**, 1527–48 (2008).
34. Case, A. J. On the Origin of Superoxide Dismutase: An Evolutionary Perspective of Superoxide-Mediated Redox Signaling. *Antioxidants (Basel, Switzerland)* **6** (2017).
35. Miller, A.-F. F. Superoxide dismutases: Ancient enzymes and new insights. *FEBS Lett.* **586**, 585–595 (2012).
36. Benzie, I. F. Evolution of antioxidant defence mechanisms. *Eur. J. Nutr.* **39**, 53–61 (2000).
37. Cao, Y. *et al.* Structural, Evolutionary, and Functional Analysis of the Class III Peroxidase Gene Family in Chinese Pear (*Pyrus bretschneideri*). *Front. Plant Sci.* **7**, 1–12 (2016).
38. Wang, Y., Wang, Q., Zhao, Y., Han, G. & Zhu, S. Systematic analysis of maize class III peroxidase gene family reveals a conserved subfamily involved in abiotic stress response. *Gene* **566**, 95–108 (2015).
39. Passardi, F., Cosio, C., Penel, C. & Dunand, C. Peroxidases have more functions than a Swiss army knife. *Plant Cell Rep.* **24**, 255–265 (2005).
40. Passardi, F., Longet, D., Penel, C. & Dunand, C. The class III peroxidase multigenic family in rice and its evolution in land plants. *Phytochemistry* **65**, 1879–1893 (2004).
41. Mathé, C., Barre, A., Jourda, C. & Dunand, C. Evolution and expression of class III peroxidases. *Arch. Biochem. Biophys.* **500**, 58–65 (2010).
42. Kupriyanova, E. V., Mamoshina, P. O. & Ezhova, T. A. Evolutionary divergence of Arabidopsis thaliana classical peroxidases. *Biochem.* **80**, 1362–1372 (2015).
43. Duroux, L. & Welinder, K. G. The Peroxidase Gene Family in Plants: A Phylogenetic Overview. *J. Mol. Evol.* **57**, 397–407 (2003).
44. Passardi, F. *et al.* Prokaryotic origins of the non-animal peroxidase superfamily and organelle-mediated transmission to eukaryotes. *Genomics* **89**, 567–579 (2007).
45. Shigeto, J. & Tsutsumi, Y. Diverse functions and reactions of class III peroxidases. *New Phytol.* **209**, 1395–1402 (2016).
46. Petrov, V. D. & Van Breusegem, F. Hydrogen peroxide—a central hub for information flow in plant cells. *AoB Plants* **2012**, pls014 (2012).
47. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
48. Bakalovic, N. *et al.* PeroxiBase: A class III plant peroxidase database. *Phytochemistry* **67**, 534–539 (2006).
49. Lamesch, P. *et al.* The Arabidopsis Information Resource (TAIR): Improved gene annotation and new tools. *Nucleic Acids Res.* **40**, 1–9 (2012).
50. Szklarczyk, D. *et al.* STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, 447–452 (2014).
51. Szklarczyk, D. *et al.* STITCH 5: Augmenting protein–chemical interaction networks with tissue and affinity data. *Nucleic Acids Res.* **44**, D380–D384 (2016).
52. Castro, M. A. A., Wang, X., Fletcher, M. N. C., Meyer, K. B. & Markowitz, F. RedeR: R/Bioconductor package for representing modular structures, nested networks and multiple levels of hierarchical associations. *Genome Biol.* **13**, R29 (2012).
53. Dalmolin, R. J. S. & Castro, M. A. A. Geneplast: evolutionary rooting and plasticity inference. (2015).
54. Mirkin, B. G., Fenner, T. I., Galperin, M. Y. & Koonin, E. V. Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evol. Biol.* **3**, 34 (2003).
55. Castro, M. A. A., Dalmolin, R. J. S., Moreira, J. C. F., Mombach, J. C. M. & de Almeida, R. M. C. Evolutionary origins of human apoptosis and genome-stability gene networks. *Nucleic Acids Res.* **36**, 6269–6283 (2008).
56. UniProt: a hub for protein information. *Nucleic Acids Res.* **43**, D204–D212 (2015).
57. Kinsella, R. J. *et al.* Ensembl BioMart: A hub for data retrieval across taxonomic space. *Database* **2011**, 1–9 (2011).

Acknowledgements

We acknowledge TAIR, Gene Ontology Consortium, PeroxiBase, STRING, STITCH, and Bioconductor databases for providing public access to their data and tools. The authors would like to thank the NPAD/UFRN for computational resources. This work have been financed by the governmental Brazilian agency National Council for Scientific and Technological Development (CNPq, Portuguese: *Conselho Nacional de Desenvolvimento Científico e Tecnológico - website: cnpq.br*), grant 444856/2014-5. The scholarships were financed by governmental Brazilian agency Coordination for the Improvement of Higher Education Personnel (CAPES - Portuguese: *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - website: www.capes.gov.br*), grant 3381/2013.

Author contributions

R.J.S.D. conceived the project. R.A.C.O., J.G.S.L., J.P.M.S.L., M.A.B.P. gathered and curated the data. R.A.C.O., R.V.M.A., D.O.I., and A.S.A. performed the analysis. R.V.M.A., D.O.I. and A.S.A. organized Supplementary Material. R.A.C.O. and R.J.S.D. wrote the manuscript. All authors revised and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-52299-y>.

Correspondence and requests for materials should be addressed to R.J.S.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019