BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 10(3)

Modeling and phylogenetic analysis of cytosolic ascorbate peroxidase (OsAPX1) from rice reveal signature motifs that may play a role in stress tolerance

Saurabh Pandey[1, 2], Yogesh Kumar Negi[3], Subramanyam Chinreddy[1], Krishnamurthy Sathelly[1], Sandeep Arora[4] & Tanushri Kaul[1]*

[1]Plant Molecular Biology Laboratory, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110067, India; [2]Uttarakhand Technical University, Dehradun – 248007, India; [3]Department of Basic Sciences, College of Forestry & Hill Agriculture, Uttarakhand University of Horticulture & Forestry, Ranichauri 249199, India; [4]Department of Molecular Biology and Genetic Engineering, College of Basic Science and Humanities, Govind VallabhPant Agriculture University and Technology, Pantnagar, Uttarakahand-263145, India; Tanushri Kaul - Email: tanushrikaul@gmail.com; Phone: +91-11-26742962; Fax: +91-11-26742316; *Corresponding author

Received January 26, 2014; Accepted January 31, 2014; Published March 19, 2014

Abstract:
Ascorbate peroxidase (APX) is a crucial, haeme-containing enzyme of the ascorbate glutathione cycle that detoxifies reactive oxygen species in plants by catalyzing the conversion of hydrogen peroxide to water using ascorbate as a specific electron donor. Different APX isoforms are present in discrete subcellular compartments in rice and their expression is stress regulated. We revealed the homology model of OsAPX1 protein using the crystal structure of soybean GmAPX1 (PDB ID: 2XIF) as template by Modeller 9.12. The resultant OsAPX1 model structure was refined by PROCHECK, ProSA, Verify3D and RMSD that indicated the model structure is reliable with 83 % amino acid sequence identity with template, RMSD (1.4 Å), Verify3D (86.06 %), Zscores (-8.44) and Ramachandran plot analysis showed that conformations for 94.6% of amino acid residues are within the most favoured regions. Investigation revealed two conserved signatures for haeme ligand binding and peroxidase activity in the alpha helical region that may play a significant role during stress.

Keywords: Haeme ligand, homology modeling, ascorbate peroxidase, *Oryza sativa.*

Background:
Adverse effects on plant productivity during environmental stresses are primarily due to the production of reactive oxygen species (ROS). Plants have evolved efficient antioxidant systems to scavenge cellular ROS by an interconnected network of antioxidants and ROS-scavenging enzymes **[1].** In plants, the ascorbate-glutathione cycle that exists in every cell is a coordinated chain of biochemical reactions that performs a crucial role in H_2O_2 homeostasis. One of crucial enzyme of this cycle is ascorbate peroxidase (APX; EC 1.11.1.11) a haeme - containing protein that plays a central role in scavenging H_2O_2 to protect higher plants from oxidative stress. Under physiological conditions, it is responsible for the rapid reduction of H_2O_2 to water at the expense of two ascorbate (AsA) molecules as specific electron donors. Several isoforms of APX have been reported in *Arabidopsis thaliana* and *Oryza sativa* **[2].** APX isoenzymes are distributed in various cellular compartments, i.e. stromal and thylakoid membrane of

chloroplasts, microbodies including peroxisomes and glyoxysomes, cytosol, mitochondria **[3].** The cytosolic APXs are highly responsive to various types of stresses, even though the cytosol is not the dominant source of reactive oxygen species. The cytosolic APXs are central regulators of cellular ROS levels and are essential for regulating cellular redox states in *Oryza sativa* **[4].**

Structural studies on bio-molecules have changed our perception of the biological world in the last few years. Homology modeling is one such technique that has found wide appreciation, which aims at predicting the 3D structure of bio-molecules, relying largely on resources such as pattern/function and sequence. There is great concern regarding cytosolic APX genes in plants for enhancing tolerance to abiotic stresses. Rice being a staple food crop throughout the globe, thus, revealing the three dimensional (3D) structure of cytosolic APXs from rice would provide insights in to their functional mechanism during stress. However, 3D structure of cytosolic APX1 from *Oryza sativa* remains unknown. In the present study, we generate the 3D structure of the cytosolic OsAPX1 (EC 1.11.1.11) from *Oryza sativa* based on the available cytosolic soybean APX (2XIF) structural homologue as template. The model structure OsAPX1 was validated with standard parameters (PROCHECK, PROSAII, Verify3D, RMSD). This study could prove useful in functional characterization of plant ascorbate peroxidases in rice in response to stress.

Methodology:

Sequence search

The FASTA sequence of cytosolic APX1 (A2XFC7.1) from *Oryza sativa* was retrieved from the NCBI database (National Center for Biotechnology Information). Comparative modeling usually starts by searching the PDB of known protein structures using the target sequence as the query **[5].** This search is generally done by comparing the target sequence with the sequence of each of the structures in the database. The target sequence was searched for similar sequence using the BLAST (Basic Local Alignment Search Tool) **[6]** against Protein Database (PDB). The BLAST results yielded X-ray structure of 2XIF cytosolic APX1 from soybean with 83% similarity to our target protein (OsAPX1).

Comparative modeling

The theoretical structure of OsAPX1 from *Oryza sativa* was generated using Modeller 9.12 by comparative modeling of protein structure prediction. MODELLER implements comparative protein structure modeling by satisfaction of spatial restraints. The program was designed to use as many different types of information about the target sequence as possible.

Model validation of OsAPX1

The model was evaluated on the basis of geometrical and stereo-chemical constraints using PROCHECK, ProSA-Web **[7]** and Verify 3D **[8].**

RMSD

Root Mean Squared Deviation (RMSD) generally represents the distance between two objects. In a structural sense, this value indicates the degree to which two three dimensional structures are similar. The lower the value, the more similar the structures are. The RMSD value between the template 2XIF and our model structure was calculated using MOE.

Phylogenetic analysis

Phylogenetic analysis of the sequences was done by Molecular Evolutionary Genetic Analysis (MEGA) software (version 4.0.02) **[9],** using UPGMA method. Each node was tested using the bootstrap approach by taking 1,000 replicates.

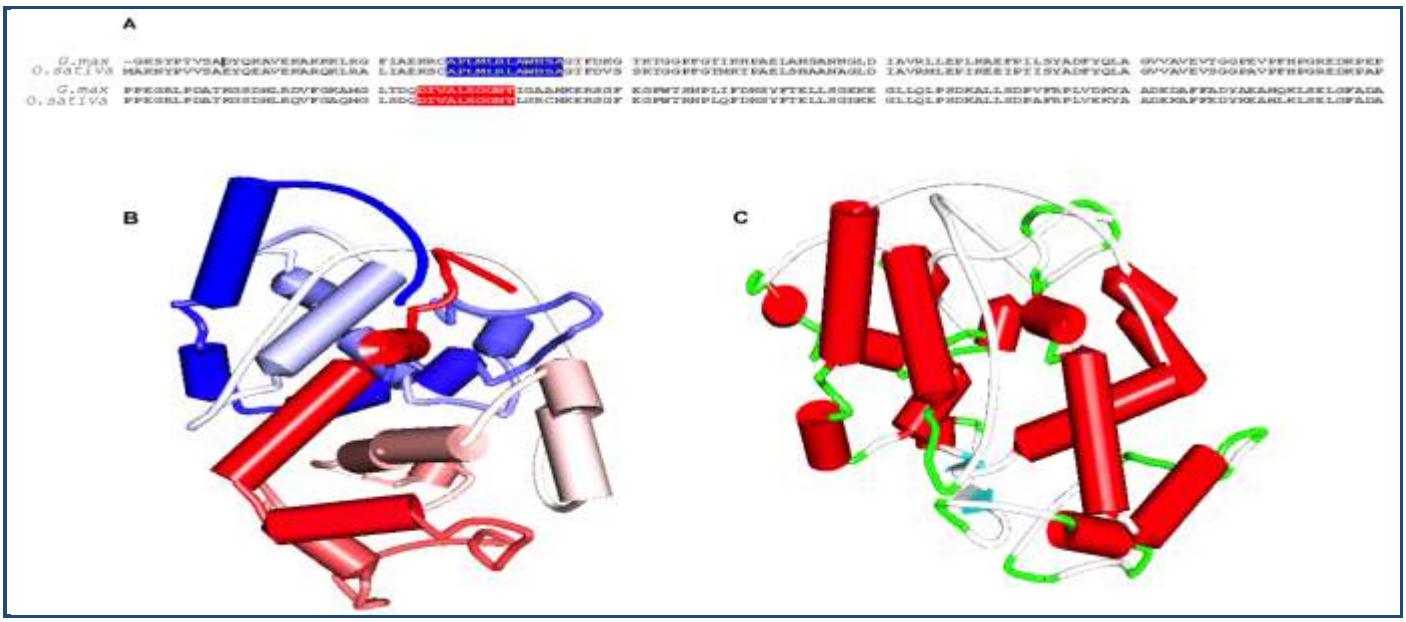

Figure 1: A) Alignment of deduced amino acid sequences of cytosolic ascorbate peroxidases from *Glycine max* (*G. max*; template) and *Oryza sativa* (*O. sativa*; target) revealing the signature motifs for peroxidase activity and haeme ligand binding shown in blue and red, respectively; **B)** Cartoon structure of OsAPX1 model showing its N- and C- terminals in blue and red, respectively; **C)** Alpha helical model of OsAPX1 showing ten helices and two sheets.

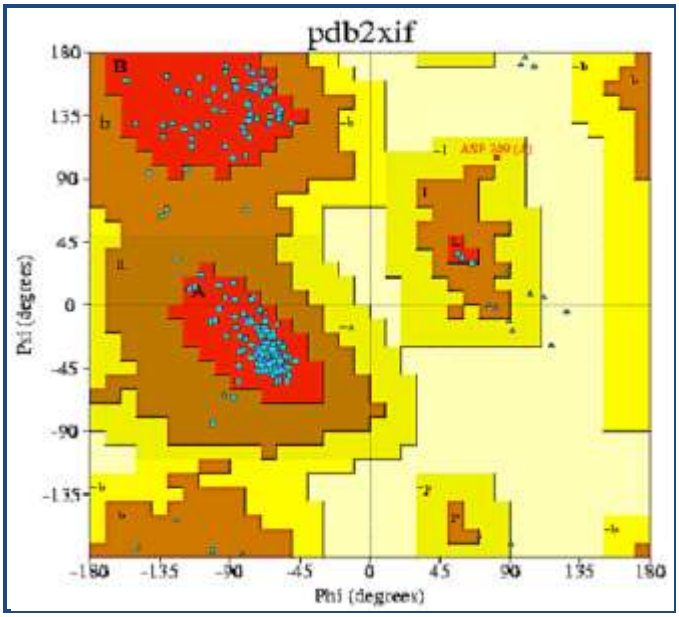

Figure 2: Plot for OsAPX1 designed by PROCHECK program. All residues of OsAPX1 are in most favoured region.

Results & Discussion:

Comparative modeling of cytosolic APX1

The rice protein sequence comprises of 250 amino acid residues. The homology search using PDB blast analysis revealed 83 % sequence identity to cytosolic APX from soybean (PDB ID: 2XIF) with an e-value of 6e-153. ScanProsite server identified the fragment APLMLRLAWHSA as the plant peroxidase signature motif (33-44 residues) as a consensus pattern in both rice and soybean APXs. Sequence alignment revealed that haeme ligand interacting signature motif IVALSGAHTL was also conserved in both target and template APX sequences **(Figure 1).** Further we developed five model structures for cytosolic OsAPX1 using Modeller 9.12. The model with the lowest DOPE (Discrete Optimized Protein Energy, a statistical potential used to assess homology models) score of -33107.04 was considered to be thermodynamically stable and chosen for further refinement and validation. This was later visualized by Accelrys Discovery Studio Version 2.5.

Validation of OsAPX1 structure

The stereochemical quality and accuracy of the predicted model was evaluated using Modeller 9.12, which is based on Ramachandran plot calculation. In general, a score close to 100 % implies good stereochemical quality of the models. The reliability of the target proteins was examined by torsion angles Ø and Ψ and a percentage quality measurement of the protein structure was used, in which four sorts of occupancies were called core, allowed, generously allowed and disallowed regions respectively **(Figure 2).** The Modeller 9.12 generated model revealed 94.6 % residues falling in most favoured region, 4.9% residues in additionally allowed region, and 0.5% residues in generously allowed region with no residues in the disallowed region of the Ramachandran plot.

ProSA-Web analysis of the model revealed a Z-score value of target protein. The Z-score indicates overall model quality and measures the deviation of the total energy of the structure with respect to an energy distribution derived from random conformations. In order to facilitate interpretation of the Z-score of the specified protein, its particular value was displayed in a plot that contains the Z scores of all experimentally determined protein chains in OsAPX1 PDB. Groups of structures from different sources (X-ray, NMR) are distinguished by different colours (NMR in dark blue and X ray in light blue). This plot can be used to check whether the Z-score of the protein in question is within the range of scores typically found for proteins of similar size belonging to one of these groups. It can be seen in **(Figure 3A)** that Z-score value (-8.04) of the target model of cytosolic OsAPX1 is located within the space of proteins determined by X ray crystallography and NMR. This value was extremely close to the value of template GmAPX (-8.6), which suggested that the obtained model was reliable and very close to experimentally determined structures **(Figure 3B).** Verify3D showed 86.06% of the residues had a score greater than 0.2 that corresponded to the quality of the OsAPX1 model that was reliable and acceptable. The degree of structural similarity was measured by RSMD performed between equivalent atom pairs. RSMD analysis of the OsAPX1 model was measured from its template (2XIF) from soybean using MOE software. The Cα RMSD and backbone RSMD deviation for the OsAPX1 model and the GmAPX template (2XIF) crystal structure were 1.06 Å, and 1.04 Å, respectively and over all RMSD was 1.57 Å **(Figure 4).** Thus, the OsAPX1 model generated by Modeller 9.12 was confirmed to be reliable and accurate.

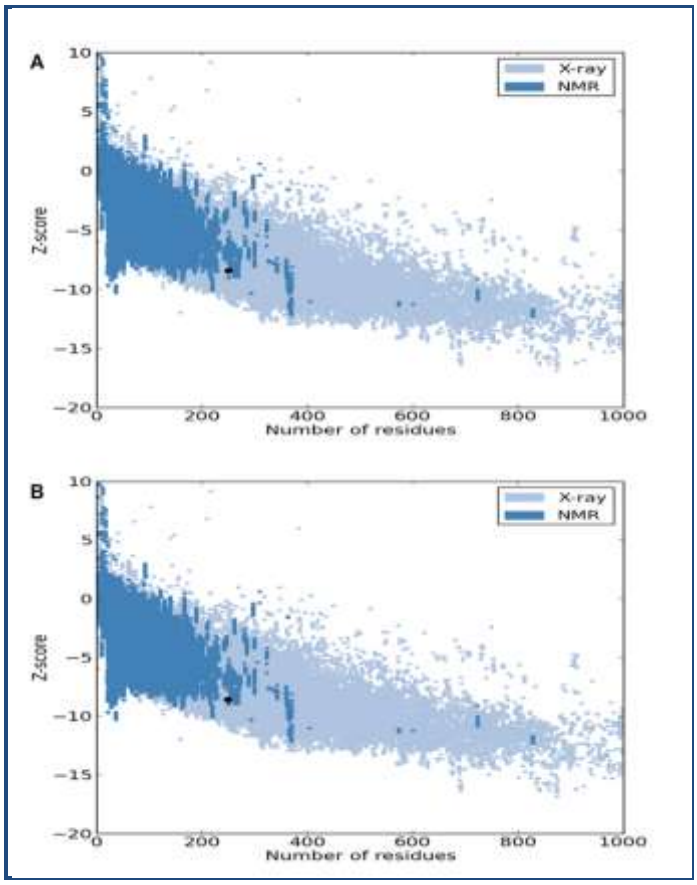

Figure 3: Validation of OsAPX1 structure model by ProSAII tools. The Z-score values of **A)** OsAPX1 (Target) and **B)** GmAPX1 (Template) proteins determined by NMR (represented in dark blue) and X-ray (represented in light blue). The two black dots represent Z-score value of template and target.

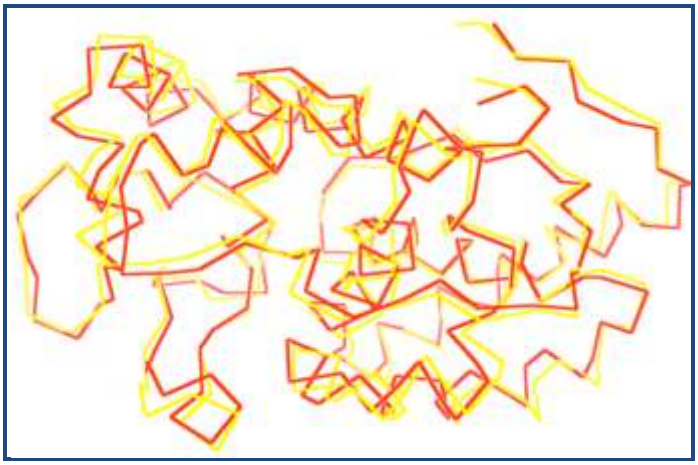

Figure 4: Superposition of Cα backbone of OsAPX1 (target) and GmAPX (template PDBID: 2XIF) represented by red and yellow colours, respectively.

Comparative structural alignment of OsAPX1 protein with other plant homologues

Comparative alignment of OsAPX1 with its homologous protein sequences revealed that helices (10) and strands (2) were conserved in all the selected plant species **(Figure 5).** High sequence similarity was observed at both the C- and N- termini. The haeme ligand binding and peroxidase signature motifs were present in right handed alpha helix **(Figure 4).** Our analysis revealed that alpha helix might be involved in haeme ligand interactions. In addition, it revealed that haeme ligand signature or consensus pattern ([D/E]IVALSG[G/A]T[L/I]G) was conserved in dicots and monocots (as rice being monocot and soybean dicot) with 83% similarity. Thus, the model provides insight into the molecular function of the haeme binding residues and peroxidase activity of OsAPX1 in response to stress.

Phylogenetic analysis

The APX proteins were highly conserved across all higher plants and encoded by a multigene family. Duplication of genes and their subsequent divergence were central to multiplicity of the APX gene family. To establish and evolutionary relationship of OsAPX1 (used in this study) with other plant orthologs, we selected protein sequences available in NCBI database for cytosolic APXs from monocots (10) and dicots (9) for phylogenetic analysis. Phylogenetic analysis showed a clear demarcation of plant cytosolic APXs into two prominent clusters; cluster A and B. Cluster A comprised of sequences from monocots *(Aeluropus littoralis, Eleusine coracana, Setaria italica, Zea mays, Sorghum bicolour, Brachypodium distanchyon, Hordeum vulgare, Oryza sativa, Triticum aestivum*), whereas cluster B included sequences from dicots (*Theobroma cacao, Gossypium hirsutum, Medicago truncatula, Glycine max, Bruguieran gymnorhiza, Brassica rapa, Brassica oleracea, Nicotiana tabacum, Solanum lycopersicum, Capsicum annuum*) **(Figure 6).** Cluster A has two subgroups a and b wherein; subgroup a included (*Eleusine coracana*, *Setaria italica*, *Zea mays*) and subgroup b contained (*Triticum aestivum*, *Hordeum vulgare, Oryza sativa*). On the other hand cluster B comprised of subgroups c, d and e, wherein; subgroup c contained (*Theobroma cacao, Gossypium hirsutum*), subgroup d contained (*Medicago truncatula* and *Glycine max*) and subgroup e contained (*Brassica rapa, Brassica oleracea, Nicotiana tabacum, Solanum lycopersicum* and *Capsicum annuum*). However, *H. vulgare* and *B. distanchyon* showed highest sequence similarity to OsAPX1. All the major clusters gave bootstrap values higher than 60. The tree showed distinct crop-specific clustering of sequences, revealing clear crop-specific differences.

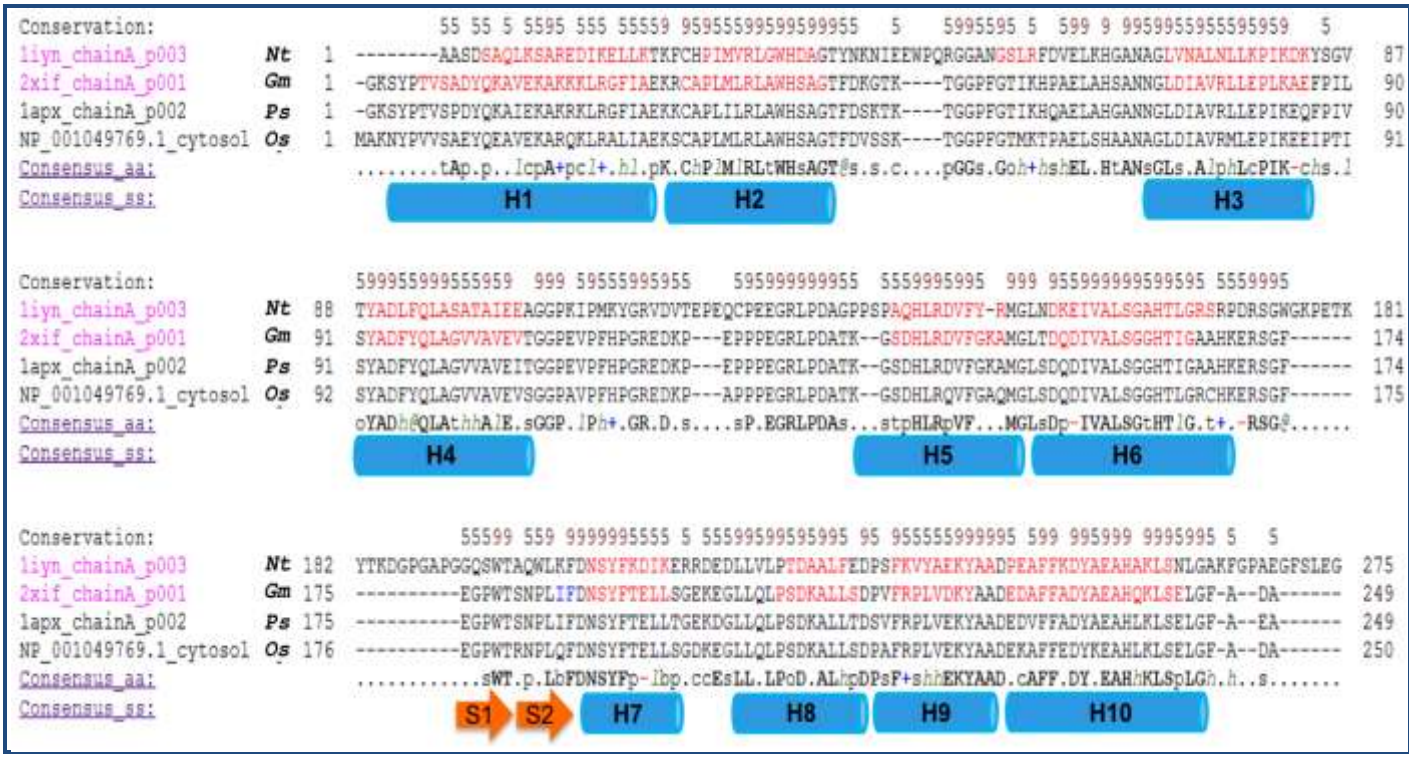

Figure 5: Comparative alignment of deduced amino acid sequences of OsAPX1 and other plant homologues showing peroxidase signature motif (H 1 & 2) in blue and the haeme ligand signature motif (H 6) in black.

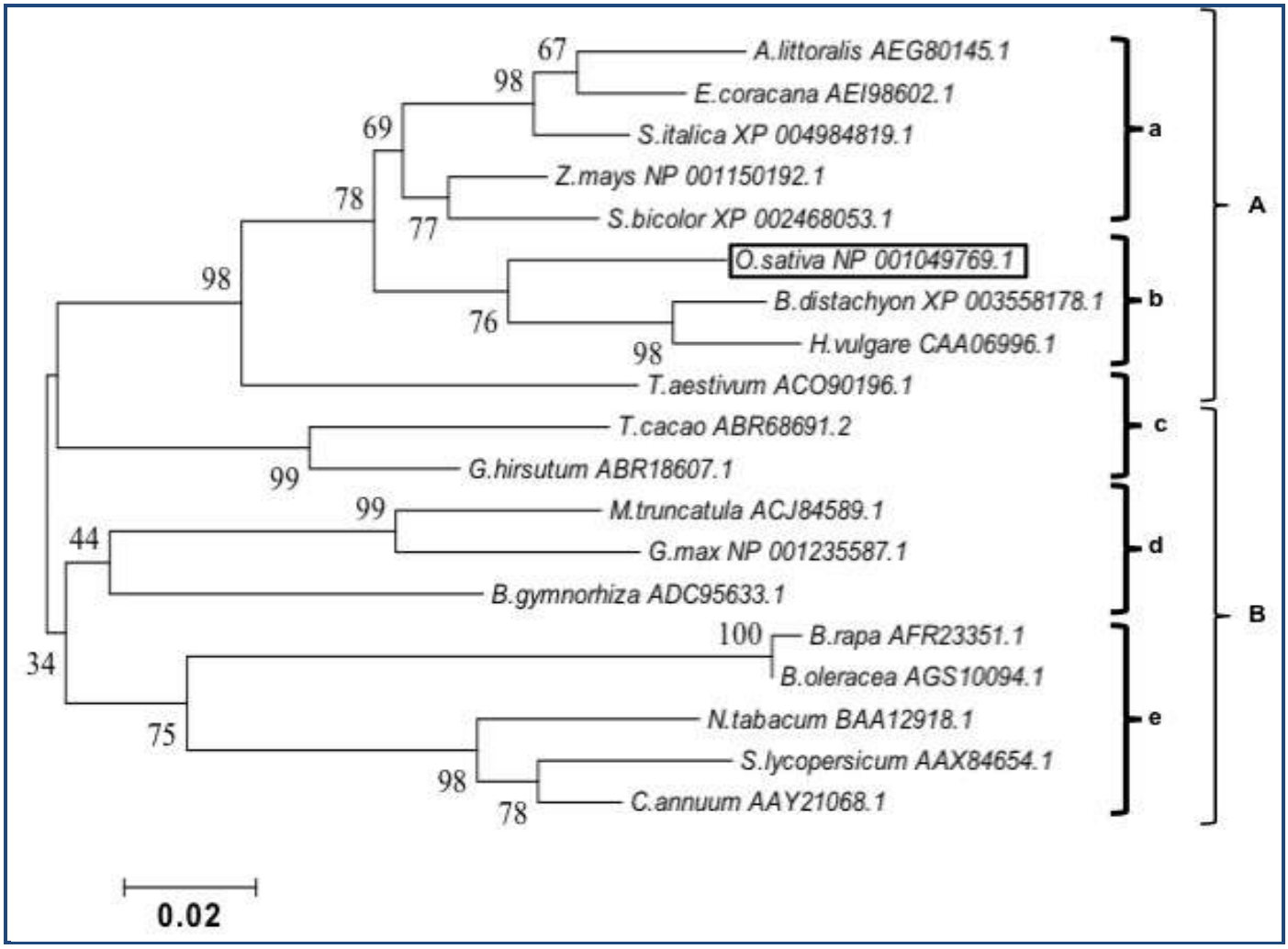

Figure 6: Phylogenetic analysis showing similarity of rice cytosolic ascorbate peroxidase (OsAPX1) with other cytosolic ascorbate peroxidases from dicots and monocots.

Conclusion:

To our knowledge this is the first report regarding rice cytosolic APX1 homology modeling. Ascorbate peroxidase is a key enzyme for regulation of ROS levels in different subcellular compartments that plays a significant role in conferring stress tolerance in plants. The model provides insight into the molecular functions of haeme ligand binding and peroxidase signature motifs during stress resistance in plants. Hence, the model structure of OsAPX1 shall prove to be crucial in comprehending its molecular function in response to stress.

Acknowledgement:

Authors like to appreciate and thank the Department of Biotechnology for funding this research.

References:

[1] Ara N *et al. Int J Mol Sci.* 2013 **14**: 24008 [PMID: 24336062]

[2] Caverzan A *et al. Genet Mol Biol.* 2012 **35:** 1011 [PMID: 23412747]

[3] Pandey S *et al. J Bioengineer & Biomedical Sci* **1**: 103. doi:10.4172/2155-9538.1000103

[4] Xu Q *et al. PLoS One.* 2013 **8:** e81640 [PMID: 24312331]

[5] Westbrook J *et al. Nucleic Acids Res.* 2002 **30**: 245 [PMID: 11752306]

[6] Altschul SF *et al. Nucleic Acids Res.* 1997 **25**: 3389 [PMID: 9254694]

[7] Wiederstein M & Sippl MJ, *Nucleic Acids Res*. 2007 **35**: W407 [PMID: 17517781]

[8] Eisenberg D *et al*. *Methods Enzymol*. 1997 **277**: 396 [PMID: 9379925]

[9] Tamura K *et al. Mol Biol Evol.* 2007 **24**: 1596 [PMID: 17488738]

Edited by P Kangueane

Citation: **Pandey *et al.*** Bioinformation 10(3): 119-123 (2014)