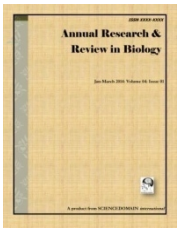

Annual Research & Review in Biology
4(24): 3810-3815, 2014

SCIENCEDOMAIN *international*
www.sciencedomain.org

Phylogenetic Trends of Plant Glutathione peroxidases Revealed by Kohonen Maps (SOM's)

Sayak Ganguli[1*] and Abhijit Datta[2]

[1]DBT Centre for Bioinformatics, Presidency University, Kolkata, India.
[2]Department of Botany, Jhargram Raj College, West Bengal, India.

Authors' contributions

This work was carried out in collaboration between both authors. Author SG designed the study and performed the analyses while author AD was involved in overall quality assurance and validating the results. Both the authors approve of the final manuscript.

Original Research Article

Received 7th April 2014
Accepted 6th June 2014
Published 8th July 2014

ABSTRACT

Aims: Glutathione peroxidases in plants have diverse roles in the maintenance of redox homeostasis in conditions of stress. The studies involving this group of enzymes have clearly demonstrated that they are near homologues to animal phospholipid hydroperoxide glutathione peroxidases with cysteine replacing selenocysteine at the sites of interaction for ligands. Phylogenetic insights of the group should provide us with indications regarding the transition of the residues needed for interactions throughout the lineage of plants for this enzyme right from the aquatic members through land plants and to the highest evolved plant groups.
Study Design: Glutathione peroxidase gene sets along with other peroxidase gene sets were retrieved from the existing databases and self organizing maps were generated.
Place and Duration of the Study: The entire study was performed at the DBT Centre for Bioinformatics, Presidency University, Kolkata for a period of June 2013 to January 2014.
Methodology: A comparative clustering was performed using self organizing maps – A technique for comparative unsupervised learning and standard neighbor joining and UPGMA methods of tree generation to identify and delineate clusters of glutathione peroxidase genes.
Results: Homology in clustering was observed when the phylogentic tree and the self organizing maps were compared. Specific sister groups were found to occupy unique

**Corresponding author: Email: sayakbif@yahoo.com;*

areas as part of the self organizing map.
Conclusion: Observation and results indicate that GPX evolution follows the trend of plant transitory evolution.

Keywords: Kohonen Maps; glutathione peroxidase; neighbor joining; UPGMA.

1. INTRODUCTION

Several antioxidant defense pathways have been developed by the aerobic organism genome to counter the damage induced by reactive oxygen species. These chemical entities are generated during mitochondrial respiration or cytosolic metabolism via an incomplete reduction of oxygen molecules. These molecules post production play important roles in various signaling cascades [1].

Enzyme systems such as catalases and peroxidases have been identified and characterized from numerous species under conditions of stress. Berry et al. [2,3] have reported extensively how glutathione acts as an electron donor during the reduction of H_2O_2 and organic hydroperoxides in a selenium dependent mechanism. Mammalian GpX families have been well characterized over the years and our current understanding limits the total number of GPX families to eight. GpX 7 and GpX 8 are vertebrate specific, lack oligomerization loop and are the least characterized [4].

Plant GpXs differ from the animal counterparts primarily by the absence of selenocysteine [5]. Apart from this they reduce peroxide with higher efficiency by using the thioredoxin system rather than glutathione. An interesting fact is that some plant GpXs exhibit striking similarities with animal phospholipid hydroperoxide glutathione peroxidases [6,8-10]. Evolutionary explorations of plant glutathione peroxidases are very limited and the available ones are attempt mainly to explain the reasons for the loss of selenocysteine from the plant group. However, proper delineation of the sequence clans and exploration of sequence structure relationships are unexplored. This analyses attempts to explore the phylogenetic relationships of glutathione peroxidases through a comparative analyses pipeline involving neural network based calculations leading to the generation of SOM's and neighbor joining algorithm for phylogenetic tree generation.

2. MATERIALS AND METHODS

2.1 Data Mining and Curation

Keyword based search strings were employed for searching the available data repositories for the desired peroxidase sequences belonging to plants. Boolean operators were used to refine the search and NCBI – Genpept [7], Uniprot and PIR databases were used for retrieval of the sequences. A basic in house PERL script was then employed to identify redundancies and these were subsequently removed. Thus the total dataset for the training was 1300 sequences while the test set was of 400 sequences. [Files containing accession numbers provided as supplementary information]

2.2 Generation of Kohonen Maps

The SOM consists of a set U of units or neurons arranged on a regular grid. Each unit I ∈ U is assigned a prototype vector m_i = [m_i 1, m_i 2... m in] where n is the number of dimensions of the input space. The neurons are connected to adjacent neurons by a neighborhood relation which constitutes the topology of the map. Usually, a rectangular or hexagonal topology is used.

Self organizing maps were generated using the INCA software with default parameters. The training of the SOM was initially done with an assemblage of peroxidase sequences of plant origin which included ascorbate peroxidases, thioredoxin peroxidases, phospholipid hydroperoxidases and glutathione peroxidases. The test set was a curated dataset of 400 sequences that were obtained from available resources. Initial screening through INCA removed 5 sequences as they did not possess the requisite number of codons in their sequences. These were later checked to identify that all the five sequences were partial cds's. Finally the test set was used to generate the map. The sequence accession numbers for both the test and experimental sets are provided as supplementary files which can be obtained directly from the corresponding author.

2.3 Phylogenetic Analyses

For phylogenetic analyses the test sequences used for generating the Kohonen Map were initially aligned using UGENE tool and were then validated using SEAVIEW. Once the alignment was complete the file was used to generate a phylogenetic tree using neighbor joining algorithm available in the PHYLIP package. The robustness of the tree was checked by bootstrapping the entire calculation for 100 replicates. A total computation time of 4 hours were used for the entire tree building phase.

3. RESULTS AND DISCUSSION

5 clusters were obtained for the map generated which amounted to 23.1% of the genes and a total of 26.2% of the map was covered. The clustering can be attributed to the fact that the test set consisted of genes that were from different ecologically distinct areas – for example aquatic and land members. Apart from that woody plant groups (clusters 4 and 5) and non woody plant groups (clusters 1, 2 and 3) were also represented in the map (Fig. 1). Several workers have used SOM clustering to identify expression levels of genes in case of non plant species while the application of self organizing maps for plant data has been restricted to pest risk assessment [11] and changes in vegetation patterns [12]. INCA [13] provides default options for sequence based SOM generation.

The tree obtained (Fig. 2) through neighbor joining clearly showed four distinct sister groups and the absence of any particular root or ancestral sequence. Careful observation of the sister groups displays that clade 1 is consisted of sequences obtained from dicotyledonous land plants (Fig. 3); clade 2 consisted of members of aquatic algal genera (Fig. 3); clade 3 consisted of sequences of monocots (Fig. 3) and clade 4 consisted of members of phospholipid hydroperoxide glutathione peroxidases and certain hypothetical protein sequences having similar domain signatures to these enzymes.

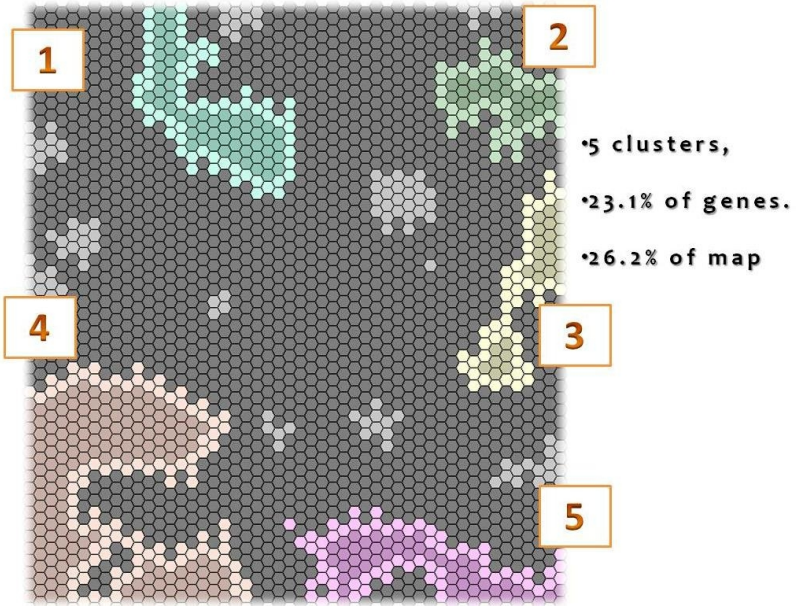

Fig. 1. Self Organizing Map of the glutathione peroxidase genes of plants

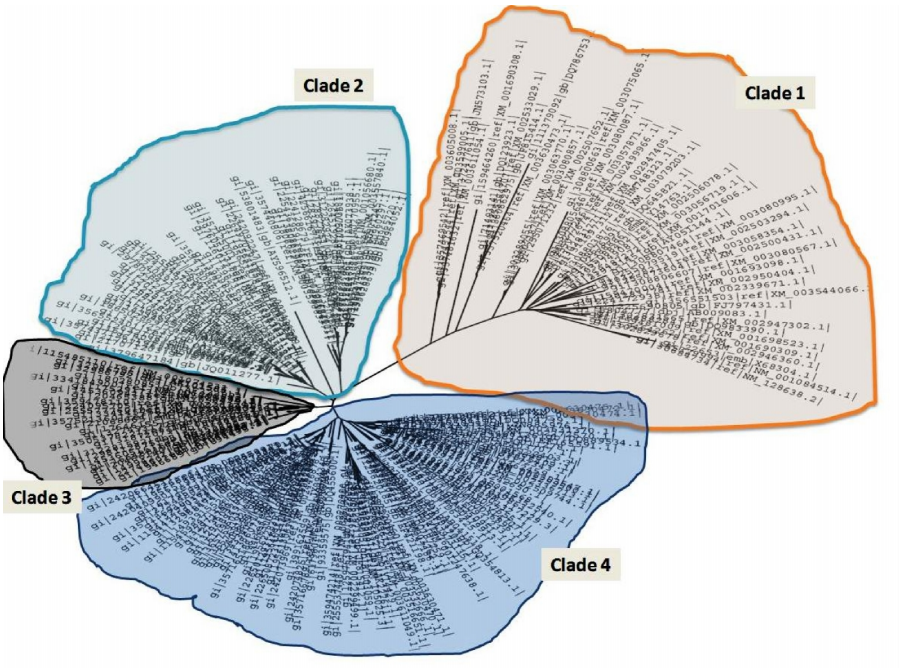

Fig. 2. Phylogenetic tree generated using nucleotide sequences of glutathione peroxidase genes of plants

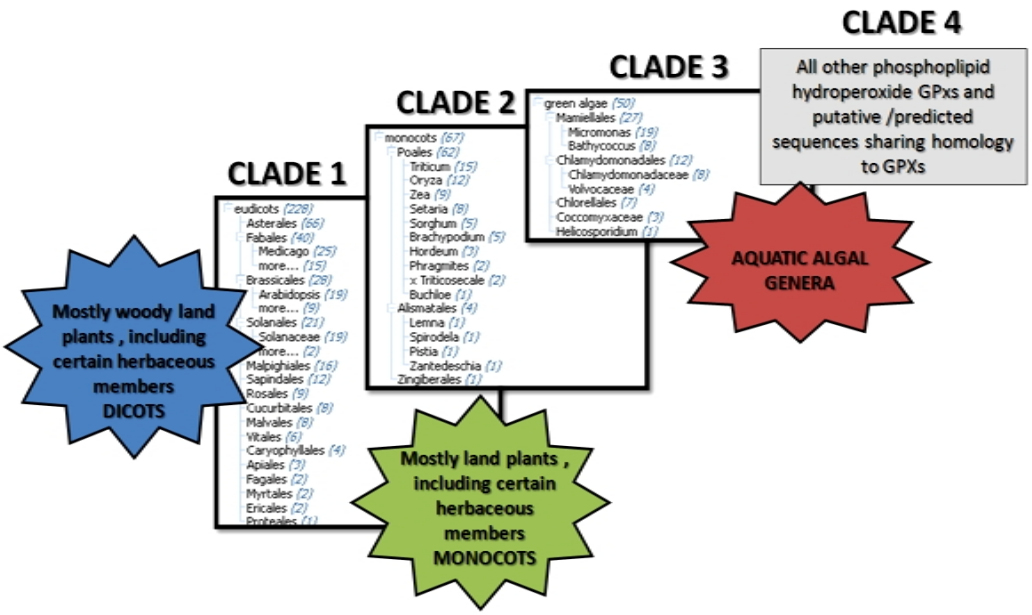

Fig. 3. Explanation of the different clades with representative members

Glutathione peroxidases, phylogenetically have been distributed in land and aquatic members since aquatic members have been reported to contain selenocysteine at the catalytic centre replaced by cysteine in case of land plant groups [6]. The tree obtained in this analysis clearly shows the demarcation in the clades (Fig. 3) following the reported tendency of evolution amongst this group of stress enzymes.

4. CONCLUSION

Thus two different clustering methods exhibited similar results depending upon the origin and environmental niche of the organisms to which the sequence belongs. These results are in conformation with the domain signatures and patterns. Clustering using Self organizing maps can now be more useful in understanding of evolutionary relationships among specific groups of genes and proteins.

CONSENT

Not applicable.

ETHICAL APPROVAL

Not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Nadaroglu H, Demir N, *In vıvo* Effects of Chlorpyrifos and Parathion Methyl On Some Oxidative Enzyme Activıties In Chickpea, Bean, Wheat, Nettle And Parsley Leaves, Fresenius Environ. Bull. 2009;18(5):647-652.
2. Berry MJ, Harney JW, Ohama T, Hatfield DL. Selenocysteine insertion or termination factors affecting UGA codon fate and complementary anticodon: Codon mutations. Nucleic Acids Res. 1994;22:3753–3759.
3. Berry MJ, Banu L, Chen YY, Mandel SJ, Kieffer JD, Harney JW, Larsen PR. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3'-untranslated region. Nature. 1991;353:273–276.
4. Gupta S, Saha S, Roy P, Basu P and Ganguli S Structural and Functional analysis of glutathione peroxidase from *Ricinus communis* L. – A computational approach, Int. Jour. Bioinformatics Res. 2010;2(1):20-30.
5. Lobanov VA, Fomenko ED, Zhang Y, Sengupta A, Hatfield DL and Gladyshev VN Evolutionary dynamics of eukaryotic selenoproteomes: large selenoproteomes may associate with aquatic life and small with terrestrial life. Genome Biology. 2007;8:R198.
6. Margis R, Dunand C, Teixeira K Felipe and Pinheiro MM. Glutathione peroxidase family – an evolutionary overview, FEBS Journal. 2008;275:3959-3970.
7. National Center for Biotechnology Information, U.S. National Library of Medicine Available: www.ncbi.nlm.nih.gov (accessed throughout the study period).
8. Obata T, Shiraiwa Y. A novel eukaryotic selenoprotein in the haptophyte alga *Emiliania huxleyi*. J. Biol. Chem. 2005;280:18462–18468.
9. Ren X, Yang L, Liu J, Su D, You D, Liu C, Zhang K, Luo G, Mu Y, Yan G, Shen J. A novel glutathione peroxidase mimic with antioxidant activity. Arch. Biochem. Biophys 2001;387:250–256.
10. Sunde RA, Evenson JK. Serine incorporation into the selenocysteine moiety of glutathione peroxidase. J. Biol. Chem. 1987;262:933–937.
11. Worner SP, Gevrey M, Eschen R, Kenis M, Paini D, Singh S, Suiter K, Watts MJ Prioritizing the risk of plant pests by clustering methods; self-organising maps, k-means and hierarchical clustering. In: Kriticos DJ, Venette RC (Eds) Advancing risk assessment models to address climate change, economics and uncertainty. Neo Biota. 2013;18:83–102.
12. Adamczyk JJ, Kurzac M, Park YS. Kruk A. - Application of a Kohonen's self-organizing map for evaluation of long-term changes in forest vegetation J. Veg. Sci. 2013;24:405–414.
13. Supek F, Vlahovicek K: Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity; BMC Bioinformatics. 2005;6:182.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=582&id=32&aid=5250