Life Science Informatics Publications

Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences

Journal Home page http://www.rjlbpcs.com/

RJLBPCS

ISSN 2454-6348

Original Research Article **DOI - 10.26479/2017.0302.07**

EVOLUTIONARY STUDY AND SEQUENCE STRUCTURE RELATIONSHIP OF FUNGAL TANNASE AND ITS SUBCELLULAR LOCALIZATION THROUGH BIOINFORMATICS

Amrita Banerjee[1], Krishanu Singha[1], Jyoti Prakash Soren[1], Arnab Sen[2], Keshab Chandra Mondal[1], Pradeep Kumar Das Mohapatra[1]*

1. Bioinformatics Infrastructure Facility Centre, Department of Microbiology, Vidyasagar University, Midnapore – 721102, West Bengal, India

2. Molecular Genetics Laboratory, Department of Botany, School of Life Sciences, University of North Bengal, Darjeeling-734013, India

ABSTRACT: Tannase (Tannin acyl hydrolase, EC 3.1.1.20), an industrially important enzyme, is predominantly produced by bacteria as well as fungi. Present study reports evolutionary relatedness among 112 fungal tannase protein sequences through superfamily/family search and protein sequences based phylogenetic analysis. Although all the selected tannase sequences belonged to alpha/beta-Hydrolase superfamily, but they were distributed in different families. Most of the *Auricularia* were related to Haloperoxidase family but maximum *Aspergillus* sp. showed relation with Hypothetical protein TT16620 family. This differentiation was also observed in phylogenetic analysis where *Auricularia* sp. and *Stereum* sp. were found in one cluster and *Aspergillus* sp. in another cluster thus indicating prominent sequence divergency. Comparison among whole protein and domain based phylogeny revealed gradual separation of related species in different blocks followed by species specific functional tannase motif identification. A total of 85 sequences were found with six common blocks among which block 2 and 4 showed relation with tannase function. Again motif 1 of block 2 (KTYWNGCSTGGREGMKQVQRF) and motif 1 of block 4 (FQERGGKMIHYHGEAD PCIPTANSVHYWQ) also showed tannase function, which could be used for species specific probe designing on PCR amplification. Subcellular localization of fungal tannase protein sequences revealed that most of the sequences were secretory in nature. Localization study will disclose the metabolic network of fungal tannase associated with mitochondria and other organelle too. Sequence and structure based analysis of *Auricularia delicata* (Acc. No. EJD47898.1) and *Aspergillus kawachii* (Acc. No. GAA89804.1) tannase protein indicated structures are more conserved than sequences and the diversity occurs at loop positions.

KEYWORDS: Fungi, tannase, phylogenetics, sequence-structure relationship, subcellular localization

***Corresponding Author: Dr. Pradeep Kumar Das Mohapatra** Ph.D.
Department of Microbiology, Vidyasagar University, Midnapore – 721102, West Bengal, India,
* Email Address: pkdmvu@gmail.com

1. INTRODUCTION

Tannase or tannin acyl hydrolase **(**EC 3.1.1.20), an inducible ester and depside bond degrading enzyme mainly produced by different microbes like bacteria, yeast and fungi. The necessity of tannase production by different organisms is to get protected from deleterious effect of plant phenolics tannins (1) and to get nutrient from the by-product gallic acid and glucose. The biotechnological application of tannase in the industrial level has gained global attention not only in the field of food stuff detannification but also in chemical and healthcare product manufacturing.Tannase is extensively used in the production of instant tea, fruit juices, alcoholic beverages, and gallic acid (2, 3). Detannification of foods and feeds by applying tannase can improve their palatability, digestibility and nutritive value (4, 2). Tannase is used in removing haze from tea beverages (5), tea cream solubilization (6) and as a biomarker for colon cancer (7). It is also used in agricultural waste and tannin polluting industrial effluents treatment (8, 9). In addition tannase plays important role in the production of antibacterial drugs (trimethoprim), propyl gallate and used as photosensitive resin in semiconductor production (10, 3). Among the different microorganisms, fungi are the most predominant tannase producers and most of the tannase industries like Biocon (India), JFC GmbH (Germany), ASA Special enzyme GmbH (Germany) and Kikkoman (Japan) supplies fungal tannase. Present study has been focused on the effect of biodiversity on fungal tannase protein sequences and structures, their subcellular localization through Bioinformatics to understand the functional conservancy and divergence pattern. It could be helpful to predict the specific primers and probes for tannase gene amplification from fungal strains and could be helpful to meet the present day demand of tannase.

2. MATERIALS AND METHODS

Sequence retrieval

A total of 112 fungal tannase protein sequences were downloaded from NCBI database and selected for evolutionary analysis.

Superfamily and family classification

Protein classification through superfamily and family analysis was done through superfamily tools on Expasy server (http://supfam.org/SUPERFAMILY/hmm.html).

Subcellular localization study

Target P V1 (11) server was used for subcellular localization study of all 112 fungal tannase protein sequences.

Conserved motif finding

All the selected protein sequences were subjected to Pfam server (http://pfam.xfam.org/search#tabview=tab1) for conserved domain finding. According to Pfam result protein domains were separated and subjected to conserved block finding through BLOCK Maker (http://blocks.fhcrc.org/blocks/blockmkr/make_blocks.html) and the results were finally subjected to MEME Suite (12) for conserved motif identification**.**

Phylogenetic tree construction

Selected sequences were aligned through ClustalX2 (13) and the multiple sequence alignment result in phylip format (.phy) was used for phylogenetic tree construction using Phylip- 3.69 (14).

Structure prediction and structural alignment

Selected sequence was subjected to structure prediction using Modeller 9.14 (15) and predicted structure was aligned with the homologous X-ray crystallographic structure using Pymol (http: //www.pymol.org.) molecular structure visualization software.

3. RESULTS AND DISCUSSION

Among the 563 downloaded fungal tannase protein sequences, 112 sequences were selected for evolutionary analysis and subcellular localization study. Among the selected sequences, 12 different genera were included like *Aspergillus spp., Auricularia sp., Botryotinia sp., Colletotrichum spp., Fusarium sp., Macrophomina sp., Neosartorya sp., Punctularia sp., Sphaerulina sp., Stereum sp., Trametes sp.* and *Verticillium sp.* The selection of sequences was based on the protein sequence diversity at least at one amino acid level. Superfamily analysis revealed that all the sequences were belonged to alpha/beta hydrolase superfamily. But differentiation was observed at the family level where 19 different families were found among 112 sequences (Table 1).Along with the inter-genus variation, intra-genus variations were also observed within *Aspergillus* spp. For three sequences of *Aspergillus fumigates* (XP_746534.1, XP_748839.1 and XP_748839.1) three different families were observed like Carbon-carbon bond hydrolase, Proline iminopeptidase-like and Serine carboxypeptidase-like respectively. Eleven *Aspergillus kawachii* were also found to distributed within four families like Haloperoxidase, Carbon-carbon bond hydrolase, Carboxylesterase and Hypothetical protein TT16620. Variation wasobserved among *Aspergillus niger* and *Aspergillus oryzae* also. Thirty one tannase protein sequences from *Colletotrichum* genus were distributed among 13 different families. On the other hand eleven sequences from genus *Auricularia delicate* were also showed relation with five different families like Halloperoxidase, Carbon-carbon bond hydrolase,

Table 1: Subcellular localization and protein family analysis of tannase protein sequences of fungi.

Sl. No.	Accession No.	Description	Subcellular localization	Family (Total sequences)
1	CCF45423.1	*Colletotrichum higginsianum*	S 3	Acetylcholinesterase-like (2)
2	EIW58147.1	*Trametes versicolor* FP-101664 SS1	S 1	
3	EJD50919.1	*Auricularia delicata* TFB-10046 SS5	_ 2	Acylamino-acid-releasing enzyme, C-terminal domain (6)
4	EQB57686.1	*Colletotrichum gloeosporioides* Cg-14	_ 5	
5	ENH74354.1	*Fusarium oxysporum* f. sp. cubense race 1	_ 2	
6	EMF16470.1	*Sphaerulina musiva* SO2202	_ 4	
7	EIM82670.1	*Stereum hirsutum* FP-91666 SS1	S 1	
8	EIM79885.1	*Stereum hirsutum* FP-91666 SS1	_ 1	
9	EKG21074.1	*Macrophomina phaseolina* MS6	S 3	Atu1826-like (4)
10	EKG21074.1	*Macrophomina phaseolina* MS6	S 3	
11	EMF08944.1	*Sphaerulina musiva* SO2202	_ 5	
12	EIW55273.1	*Trametes versicolor* FP-101664 SS1	_ 5	
13	EQB52526.1	*Colletotrichum gloeosporioides* Cg-14	S 2	Biotin biosynthesis protein BioH (8)
14	EFQ34238.1	*Colletotrichum graminicola* M1.001	S 3	
15	CCF39530.1	*Colletotrichum higginsianum*	S 1	
16	EFQ34402.1	*Colletotrichum graminicola* M1.001	S 1	
17	EKG12473.1	*Macrophomina phaseolina* MS6	_ 3	
18	EMF16470.1	*Sphaerulina musiva* SO2202	_ 4	
19	EIM91790.1	*Stereum hirsutum* FP-91666 SS1	S 1	
20	EIM87541.1	*Stereum hirsutum* FP-91666 SS1	_ 4	
21	XP_746534.1	*Aspergillus fumigatus* Af293	S 3	Carbon-carbon bond hydrolase (14)
22	XP_001389868.1	*Aspergillus niger* CBS 513.88	S 5	
23	GAA90248.1	*Aspergillus kawachii* IFO 4308	M 2	
24	GAA89737.1	*Aspergillus kawachii* IFO 4308	M 5	
25	EJD47878.1	*Auricularia delicata* TFB-10046 SS5	S 2	
26	EJD47880.1	*Auricularia delicata* TFB-10046 SS5	S 2	
27	ELA36755.1	*Colletotrichum gloeosporioides* Nara gc5	S 5	
28	ENH86659.1	*Colletotrichum orbiculare* MAFF 240422	S 1	
29	ELA31075.1	*Colletotrichum gloeosporioides* Nara gc5	_ 3	
30	ELA33543.1	*Colletotrichum gloeosporioides* Nara gc5	_ 1	
31	EIN04104.1	*Punctularia strigosozonata* HHB-11173 SS5	S 2	
32	EIN04105.1	*Punctularia strigosozonata* HHB-11173 SS5	S 2	

33	EMF08948.1	*Sphaerulina musiva* SO2202	S 4	
34	XP_003008086.1	*Verticillium alfalfae* VaMs.102	S 1	
35	GAA89804.1	*Aspergillus kawachii* IFO 4308	S 3	Carboxylesterase (6)
36	XP_001398312.1	*Aspergillus niger* CBS 513.88	S 3	
37	EJD36078.1	*Auricularia delicata* TFB-10046 SS5	_ 2	
38	EJD48968.1	*Auricularia delicata* TFB-10046 SS5	S 1	
39	EQB45005.1	*Colletotrichum gloeosporioides* Cg-14	S 1	
40	EKG20365.1	*Macrophomina phaseolina* MS6	_ 2	
41	EFQ31358.1	*Colletotrichum graminicola* M1.001	S 2	Ccg1/TafII250-interacting factor B [Cib] (1)
42	XP_001261622.1	*Neosartorya fischeri* NRRL 181	S 1	Fungal lipases (1)
43	GAA92519.1	*Aspergillus kawachii* IFO 4308	S 3	Haloperoxidase (26)
44	XP_001399173.1	*Aspergillus niger* CBS 513.88	S 4	
45	XP_001820636.1	*Aspergillus oryzae* RIB40	S 2	
46	EJD48305.1	*Auricularia delicata* TFB-10046 SS5	S 2	
47	EJD48302.1	*Auricularia delicata* TFB-10046 SS5	S 1	
48	EJD48301.1	*Auricularia delicata* TFB-10046 SS5	S 1	
49	EJD47898.1	*Auricularia delicata* TFB-10046 SS5	_ 3	
50	EJD47879.1	*Auricularia delicata* TFB-10046 SS5	_ 3	
51	EJD47885.1	*Auricularia delicata* TFB-10046 SS5	_ 3	
52	EMR89439.1	*Botryotinia fuckeliana* BcDW1	S 4	
53	EQB49046.1	*Colletotrichum gloeosporioides* Cg-14	_ 1	
54	CCF35608.1	*Colletotrichum higginsianum*	S 1	
55	ENH88685.1	*Colletotrichum orbiculare* MAFF 240422	S 4	
56	EQB46560.1	*Colletotrichum gloeosporioides* Cg-14	M 4	
57	ELA34617.1	*Colletotrichum gloeosporioides* Nara gc5	_ 2	
58	ELA28291.1	*Colletotrichum gloeosporioides* Nara gc5	M 4	
59	EFQ32873.1	*Colletotrichum graminicola* M1.001	_ 5	
60	ENH81532.1	*Colletotrichum orbiculare* MAFF 240422	_ 4	
61	CCF33605.1	*Colletotrichum higginsianum*	_ 5	
62	EMT73485.1	*Fusarium oxysporum* f. sp. cubense race 4	_ 1	
63	EMT67337.1	*Fusarium oxysporum* f. sp. cubense race 4	_ 3	
64	ENH74696.1	*Fusarium oxysporum* f. sp. cubense race 1	S 4	
65	EKG19873.1	*Macrophomina phaseolina* MS6	S 4	
67	EKG13566.1	*Macrophomina phaseolina* MS6	S 2	
68	XP_001257321.1	*Neosartorya fischeri*NRRL 181	S 3	

69	EIW55266.1	*Trametes versicolor* FP-101664 SS1	S 2	
70	EIW55273.1	*Trametes versicolor* FP-101664 SS1	_ 5	Hypothetical esterase YJL068C (1)
71	GAA88767.1	*Aspergillus kawachii* IFO 4308	S 1	Hypothetical protein TT16620 (25)
72	GAA91900.1	*Aspergillus kawachii* IFO 4308	S 4	
73	GAA92933.1	*Aspergillus kawachii* IFO 4308	S 1	
74	GAA83886.1	*Aspergillus kawachii* IFO 4308	S 3	
75	GAA82877.1	*Aspergillus kawachii* IFO 4308	S 2	
76	GAA90960.1	*Aspergillus kawachii* IFO 4308	S 3	
77	GAA92572.1	*Aspergillus kawachii* IFO 4308	_ 3	
78	XP_001388709.1	*Aspergillus niger* CBS 513.88	S 2	
79	XP_001394358.1	*Aspergillus niger* CBS 513.88	S 2	
80	XP_001402486.1	*Aspergillus niger* CBS 513.88	S 3	
81	AFB74086.1	*Aspergillus niger*	_ 2	
82	XP_001396513.1	*Aspergillus niger* CBS 513.88	S 4	
83	XP_001393409.1	*Aspergillus niger* CBS 513.88	M 2	
84	ABX89592.1	*Aspergillus niger*	M 3	
85	XP_001390411.1	*Aspergillus niger* CBS 513.88	S 2	
86	XP_001393089.1	*Aspergillus niger* CBS 513.88	S 4	
87	ABJ51876.1	*Aspergillus oryzae*	_ 2	
88	BAA09656.1	*Aspergillus oryzae*	M 5	
89	XP_001826685.1	*Aspergillus oryzae* RIB40	M 5	
90	ELA36755.1	*Colletotrichum gloeosporioides* Nara gc5	S 5	
91	ENH83447.1	*Colletotrichum orbiculare* MAFF 240422	_ 1	
92	ENH73520.1	*Fusarium oxysporum* f. sp. cubense race 1	_ 5	
93	EMT62928.1	*Fusariumoxysporum* f. sp. cubense race 4	_ 5	
94	EKG17399.1	*Macrophominaphaseolina* MS6	S 5	
95	XP_001262462.1	*Neosartorya fischeri* NRRL 181	S 3	
96	EJD50919.1	*Auricularia delicata* TFB-10046 SS5	_ 2	Hypothetical protein VC19740 (1)
97	CCF43460.1	*Colletotrichum higginsianum*	S 1	Mycobacterial antigens (2)
98	EFQ25115.1	*Colletotrichum graminicola* M1.001	_ 4	
99	EFQ27329.1	*Colletotrichum graminicola* M1.001	S 1	O-acetyltransferase (4)
100	CCF32811.1	*Colletotrichum higginsianum*	_ 4	
101	EQB47714.1	*Colletotrichum gloeosporioides* Cg-14	M 5	
102	EKG10215.1	*Macrophomina phaseolina* MS6	S 5	
103	XP_748839.1	*Aspergillus fumigatus* Af293	S 2	

104	EFQ29607.1	*Colletotrichum graminicola* M1.001	S 5	Proline iminopeptidase-like (4)
105	CCF33356.1	*Colletotrichum higginsianum*	S 1	
106	EIM82670.1	*Stereum hirsutum* FP-91666 SS1	S 1	
107	EFQ31102.1	*Colletotrichum graminicola* M1.001	_ 2	Prolyloligopeptidase, C-terminal domain (4)
108	EMT74633.1	*Fusarium oxysporum* f. sp. cubense race 4	_ 3	
109	XP_748839.1	*Aspergillus fumigatus* Af293	S 2	Serine carboxypeptidase-like (3)
110	XP_001401809.1	*Aspergillus niger* CBS 513.88	S 2	
111	ENH83447.1	*Colletotrichum orbiculare* MAFF 240422	_ 1	
112	EQB50558.1	*Colletotrichum gloeosporioides* Cg-14	S 1	Thioesterases (1)

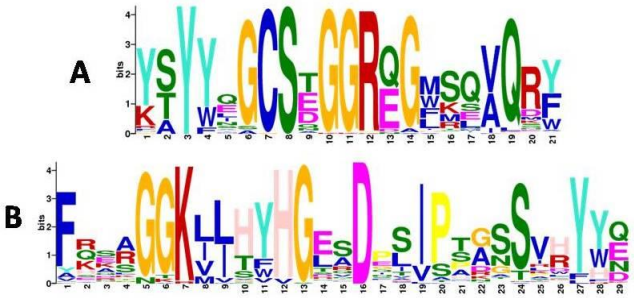
S: Secretory, M: Mitocondrial, (_): other organelle

Acylamino-acid-releasing enzyme, C-terminal domain, Carboxylesterase and an unique family of Hypothetical protein VC19740. So the family level distributions of tannase protein sequences were observed within same genus and same species also. Maximum *Aspergillus* sequences were found in Hypothetical protein TT16620 family and maximum *Auricularia* sequences were found in Haloperoxidase family. Subcellular localization study of all the individual sequences showed that among 112, 65 sequences were secretory in nature, 37 were associated with other organelle and 9 sequences were found to associate with mitochondria. Above result indicated that based on the physicochemical parameters, fungal tannase are mainly secretary and they were hydrophilic in nature. Tannase associated with mitochondria in case of *Aspergillus kawachii* IFO 4308 (GAA90248.1, GAA89737.1), *Colletotrichum gloeosporioides* Cg-14 (EQB46560.1) *Colletotrichum gloeosporioides* Nara gc5 (ELA28291.1),*Aspergillus niger* CBS 513.88 (XP_001393409.1), *Aspergillus niger*(ABX89592.1), *Aspergillus oryzae* (BAA09656.1), *Aspergillus oryzae* RIB40(XP_001826685.1), and *Colletotrichum gloeosporioides* Cg-14 (EQB47714.1) showed similarity with different families (Table 1)like Carbon-carbon bond hydrolase (GAA90248.1, GAA89737.1), Haloperoxidase (EQB46560.1, ELA28291.1), Hypothetical protein TT16620 (XP_001393409.1, ABX89592.1, BAA09656.1, XP_001826685.1), O-acetyl transferase (EQB47714.1). According to endosymbiotic theory present result indicated different combination of association between bacteria and eukaryotic ancestors of fungi.

Although the sequences showed relation with different protein families, based on their functional similarity different conserved sequence stretches were found. Among all the sequences, 85 sequences were found with six common blocks where block 2 and 4 showed relation with tannase function during BLAST analysis. Specifically motif 1 of block 2 (KTYWNGCSTGGREGMKQVQRF) and motif 1 of block 4 (FQERGGKMIHYHGEADPCIPTANSVHYWQ) showed similarity with tannase (Figure 1) during InterproScan analysis. The conserved signature sequence "GCSXGGR" of motif 1 from block 2 showed highest conservancy among all the selected sequences which was also analyzed by Banerjee et al. 2012 (16). Above motifs could be used for species specific probe

designing on PCR amplification.

Figure 1: Logo of two conserved motif among selected fungal species.



Whole protein sequence based and domain sequence based phylogenetic tree reflected the same results with respect to a bacterial tannase from *Burkholderia ambifaria* (Accession No.: EDT05970.1), used as outgroup. In both phylogram two different clads were observed were all the *Aspergillus and Auricularia* were found in either clade. Sequences from genus *Stratum* were found with *Auricularia* sp. whereas sequences from genus *Colletotrichum* have been evenly distributed between two different clades. In whole protein sequence based tree (Figure 2), *Aspergillus spp.* were found in two different groups. But in domain based tree (Figure 3), *Aspergillus* spp. were distributed in three different groups within one clad. So, domain based tree showed more precise result than whole protein based tree. So, it could be concluded that domain sequence were more conserved than whole protein sequences. Sequence based analysis of phylogenetic tree also showed variations as indicated by family level classification of tannase sequences. Diversity among orthologous proteins sharing same function, not only depends upon amino acid sequences and internal energy of tertiary protein structure, but also depends upon external minimum free energy and interaction with other physicochemical and biological environments (17, 18, 19). In this regards it may be concluded that tannase protein sequence diversity among different fungal species, was the cumulative effect of species specificity and varying interaction with external environments.

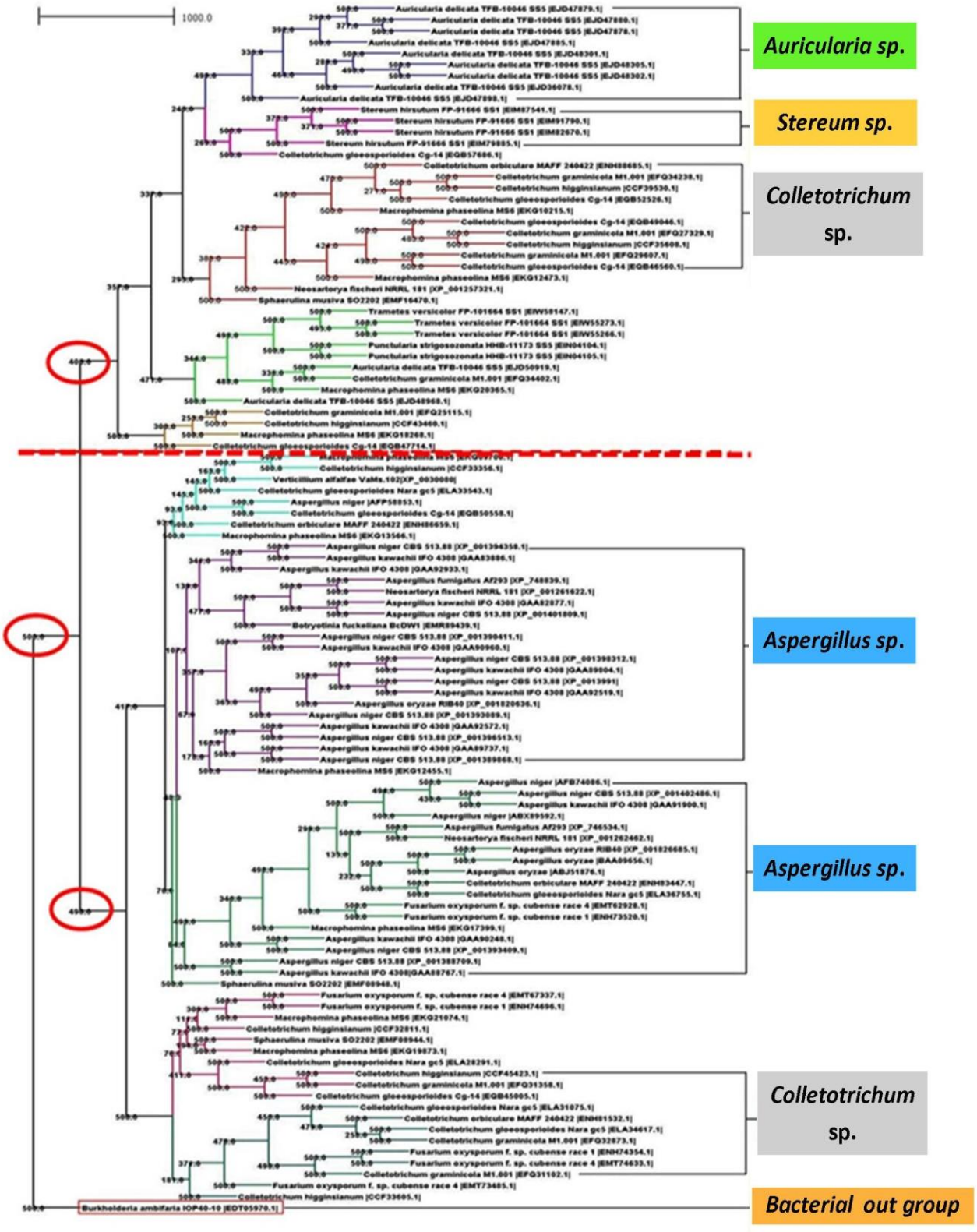


Figure 2: Whole tannase protein sequence based phylogenetic tree.

To understand the effect of sequence diversity on protein structures, two most distant sequences in domain based tree, from *Auricularia delicata* (Acc. No. EJD47898.1) and *Aspergillus kawachii* (Acc. No. GAA89804.1) were used for tertiary structure prediction based on the homologous X-ray

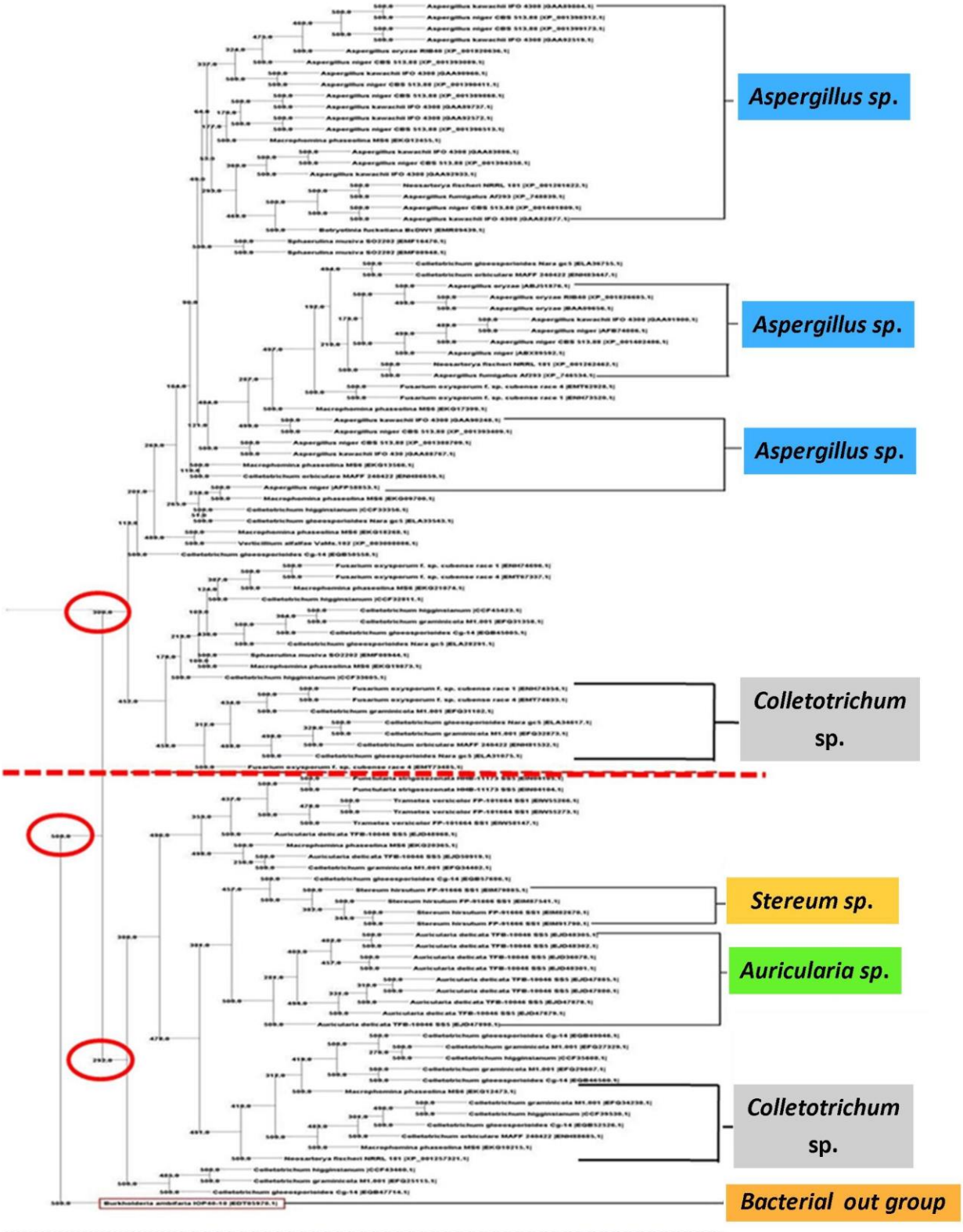


Figure 3: Tannase domain sequence based phylogenetic tree.

crystallographic structure of *Aspergillus oryzae* tannase with PDB ID: 3WMT in Protein Data Bank. From the whole protein structure based alignment, it was observed that sequence level diversity have not been reflected on structures (Figure 4A) so that very less structural distortion was noticed as yellow doted marks. Although a regular pattern of 9 β-sheet was found for both of them, diversity was

observed at α helical patterns which are differently associated apart from 9 β-sheet. During structural conservancy analysis among 9 β-sheet, the signature sequence of "GCSXG" was observed in sixth β-sheet and at the same region.

Figure 4: Alignment of tannase structures. A: *Auricularia delicata* (Acc. No. EJD47898.1) and *Aspergillus*

kawachii (Acc. No. GAA89804.1) B: Nine β sheet conservancy pattern and position of signature sequence.

A specialized structure of nine β pleated sheets was observed for both the structures which maintain the structure of functional active site constant. So, from the above result it could be understand that structures are more conserved than sequences. Whereas amino acid sequence diversity was observed at the loop position of tannase, which indicated the structural adaptation site and also reflected the environmental stability.

4. CONCLUSION

Fungi are the predominant tannase producer among all others organisms. Tannase protein sequence and structure based analysis among different fungal species could be helpful to understand the structure function relationship of a protein and adaptive nature of specific fungus with their immediate environment. Conserved motifs could be used for fungal tannase specific primer or probe designing as well as the subcellular localization study could be helpful to detect the metabolic network and horizontal gene transfer phenomenon as tannase were found to be associated with mitochondria and other organelle too.

CONFLICT OF INTEREST

The authors have no conflict of interest.

ACKNOWLEDGEMENTS

Authors are very much thankful to UGC, Govt. of India for providing fellowship to the Department of Microbiology, Vidyasagar University and also the Department of Biotechnology, Govt. of India is acknowledged gratefully for creation of BIF centre in Vidyasagar University.

REFERENCES

1. Lorenz MM, Alkhafadji L, Stringano E, Nilsson S, Mueller-Harvey I and Ud´en P. Relationship between condensed tannin structures and their ability to precipitate feed proteins in the rumen. J Sci Food Agric. 2013; 94(5):963-968.
2. Lekha PK and Lonsane BK. Production and application of tannin acyl hydrolase, state of the art. Adv Appl Microbiol. 1997; 44:216-260.
3. Das Mohapatra PK, Mondal KC and Pati BR. Production of tannase by the immobilized cells of Bacilluslicheniformis KBR6 in Ca-alginate beads. J Appl Microbiol. 2007; 102(6): 1462-1467.
4. Archambault J, Lacki K and Duvnjak Z. Conversion of catechin and tannic acid by an enzyme preparation from Trametes versicolor. Biotechnol Lett. 1996; 18(7):771-774.
5. Su E, Xia T, Gao L, Dai Q and Zhang Z. Immobilization and characterization of tannase and its haze-removing. Rev Agroquim Tecnol Aliment. 2009; 15(6): 545-552.
6. Selwal MK, Yadav A, Selwal KK, Aggarwal NK, Gupta R and Gautam SK. Tannase production by *Penicillium atramentosum* KM under SSF and its applications in wine clarification and tea cream solubilization. Braz J Microbiol. 2011; 42(1): 374-387.
7. Noguchi N, Ohashi T, Shiratori T, Narui K, Hagiwara T, Ko M, Watanabe K, Miyahara T, Taira S, Moriyasu F and Sasatsu M. Association of tannase-producing *Staphylococcus lugdunensis* with colon cancer and characterization of a novel tannase gene. J. Gastroenterol. 2007; 42(5):346-351.
8. Murugan K and Al-Sohaibani SA. Biocompatible removal of tannin and associated color from tannery effluent using the biomass and tannin acyl hydrolase (EC 3.1. 1.20) enzymes of mango industry solid waste isolate *Aspergillus candidus* MTTC 9628. Res J Microbiol. 2010; 5(4): 262-271.
9. Srivastava A and Kar R. Characterization and application of tannase produced by *Aspergillus niger* ITCC 6514.07 on pomegranate rind. Braz J Microbiol. 2009; 40(4), 782-789.
10. Hadi TA, Banerjee R and Bhattacharyya BC. Optimization of tannase biosynthesis by a newly isolated Rhizopusoryzae. Bioprocess Biosyst Eng. 1994; 11(6): 239-243.
11. Emanuelsson O, Nielsen H, Brunak S and Von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequences. J Mol Biol. 2000; 300:1005-1016.

12. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW and Noble WS. MEME SUITE, tools for motif discovery and searching. Nucleic Acids Res. 2009; 37(2): W202-W208.
13. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R and Thompson JD. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21): 2947-2948.
14. Tuimala J. A primer to phylogenetic analysis using the PHYLIP package. Cladistics. 1989; 5:164-166.
15. Eswar N, Eramian D, Webb B, Shen M-Y and Sali A. Protein Structure Modeling with modeller. Structural Proteomics. 2008; 426:145-159.
16. Banerjee A, Jana A, Pati BR, Mondal KC and Das Mohapatra PK. Characterization of tannase protein sequences of bacteria and fungi: An in silico study. The Protein Journal. 2012; 31: 306-327.
17. Anfinsen CB. Principles that govern the folding of protein chains. Science. 1973; 181: 223-230.
18. Anfinsen CB and Scheraga HA. Experimental and theoretical aspect of protein folding. Adv Protein Chemistry. 1975; 29: 205-300.
19. Li M, Huang Y and Xiao Y. Effects of external interactions on protein sequence-structure relations of beta-trefoil fold. Protein: Structure Function Bioinformatics. 2008; 72: 1161-1170.