



Journal of Experimental Biology and Agricultural Sciences

<http://www.jebas.org>

ISSN No. 2320 – 8694

EXPERIMENTAL AND COMPUTATIONAL APPROACHES FOR ASSESSMENT, BIO-DEGRADATION AND DETOXIFICATION OF PAPER INDUSTRY EFFLUENTS

Akhil Gupta*^{#1}, Apoorv Tiwari^{#2,3}, Anand Pathak³, Anamika Tripathi¹

¹Pollution Ecology Research Laboratory, Department of Botany, Hindu college Moradabad, University of Rohilkhand- 244001, India

²Department of Molecular Biology & Genetic Engineering, CBSH, GBPUA&T, Pantnagar, US Nagar 263145, Uttarakhand, India

³Sam Higginbottom University of Agriculture, Technology and Sciences (SHUATS), Allahabad, 211007, India;

⁴Department of soil sciences, Govind Ballabh Pant University of Agriculture and technology, Pantnagar US Nagar 263145 Uttarakhand, India

#Authors contributed equally.

Received – January 05, 2018; Revision – April 02, 2018; Accepted – April 24, 2018

Available Online – April 25, 2018

DOI: [http://dx.doi.org/10.18006/2018.6\(2\).425.436](http://dx.doi.org/10.18006/2018.6(2).425.436)

KEYWORDS

Pulp and paper

Effluent

Soil

In silico

Conserved domain

Motif

Phylogenetic analysis

Aspartic acid

ABSTRACT

Paper industry produced significant volumes of liquid wastes termed as ‘effluents’ which is routinely released in the approaching water bodies. In current study, author aimed to assess the effluents present in the basin of Aril River in order to understand the potential threats posed to the population due to activities like irrigation, fishing etc. Additionally, integrated approach of experimental and computational analysis was carried out to devise strategies for the bio-degradation and detoxification of the effluents to safeguard the water bodies and the environment. In this sequence, the effluents and soil samples from five different sites were collected and analyzed for different physico-chemical properties including pH, TDS, Turbidity, EC, BOD, COD, Nitrate, Phosphate, TKN, Alkalinity Chloride, Sodium Potassium, Salinity and Heavy metals every month for a period of two years from the river aril receiving effluent channel of paper industry. Present study puts forward the critical situation of different polluted sites throughout the study period as they exceeded the permissible limits. Rotated component matrix showed the source of Principal component 1&2 is effluent and Principal component 3 is earth crust. *In silico* analysis regarding the secondary and tertiary structures of proteins concerned, GRAVY score and

* Corresponding author

E-mail: akhil07123@gmail.com (Akhil Gupta)

Peer review under responsibility of Journal of Experimental Biology and Agricultural Sciences.

Production and Hosting by Horizon Publisher India [HPI]
(<http://www.horizonpublisherindia.in/>).
All rights reserved.

All the article published by Journal of Experimental Biology and Agricultural Sciences is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License Based on a work at www.jebas.org.



Prevention

Control

Motif analysis revealed molecular insights of enzymes 'in action' and study indicated that the cupridoxin family of enzymes played important role in molecular transfer reaction which can break the large recalcitrant molecules into simpler less toxic molecules. *In silico* analysis further revealed that aspartic acid based biochemistry can be crucial for detoxification of pulp and paper industry effluents.

1 Introduction

With the constantly increasing population, there arises an ever increasing demand of pulp and paper in India. In order to meet the demand and to bridge the gap between demand and supply, a number of small, medium and large scale paper mills have been operating throughout the country. The pulp and paper industry is one of the crucial industrial sectors in India, having 15th rank world-wide in paper production, generating more than 0.3 million direct and around one million indirect employments through agricultural activities. At present, there are 668 pulp and paper mills in India, in which 632 units are agro-residue and recycled fiber based units with manufacturing capacity of 7.6 million (Ministry of power, GOI, 2015).

Being highly water intensive, it consumes 100-250 m³ freshwater / ton paper and also generates the corresponding waste water 75-225 m³ wastewater / ton paper (Tewari et al., 2008). Currently the processes being adopted by this industry for the production of pulp and paper generates significant volumes of liquid wastes termed as 'effluents' which is routinely released in the approaching water bodies. However, there are regulations on how to process and treat these effluents before releasing in the environment but, there are reports that a number of such facilities have been failed to follow the guidelines. Therefore, the effluents being released into the water bodies may contain hazardous chemicals and substances that pose the great environmental pollution (Tornerio & Hanke 2016)

The value of wastewater for crop production has been recognized in arid and semi-arid regions of the world including India. In such areas, effluent has become an important source of irrigation water as well as a source of plant nutrients such as Nitrogen (N), Phosphorus (P), Potassium (K) and trace elements like Sodium (Na), Calcium (Ca) and Magnesium (Mg). Effluent irrigation can eliminate water shortage; reduce the need for chemical fertilizer and enhance the soil fertility (Mishra et al., 2013; Tripathi et al., 2014).

However, unregulated irrigation with untreated effluent poses serious public health risks, as this is a major source of heavy metals that cause accumulation in plant parts. Long-term irrigation with effluents has the potential of increasing organic

carbon and heavy metals accumulation in soil thereby multiplies the chances of their entrance in food chain and this ultimately causes high level of geo-accumulation, bio-accumulation (Chandra et al., 2010; Chopra et al., 2013). Nevertheless, certain enzymes produced by microbes found in soil and effluent of the polluted sites such as ligninases, laccases, Mnperoxidases, Xylanases and cellulases play a vital role in detoxification of pulp and paper mill effluent (Qureshi et al., 2014). The aim of current study was to understand the potential threats posed to the population due to various activities in the basin of Aril River. Integrated approach of experimental and computational analysis was carried out to devise strategies for the bio-degradation and detoxification of the effluents to safeguard the water bodies and the environment.

2 Material and Methods

2.1 Study sites and Physico-chemical characteristics

Five sampling sites (1, 2, 3, 4 & 5) were located around the paper mill (Shakumbhari Straw Products limited) situated on Moradabad - Aligarh Highway, NH 93, Sites were selected on the basis of distance of one km along the effluent channel. India. Various Physico-chemical characteristics such as EC, pH, color, Turbidity, TDS, DO, BOD, COD, Na, K, Chlorides, Alkalinity, Nitrate, TKN, Salinity, Phosphate, Cu, Zn, Cd and Cr were taken as parameters of samples. The temperature of the sample was determined by mercury thermometer. The pH of the sample was measured calibrating digital pH meter (Systronics MK6) with buffer solution of pH 4 and 9. Turbidity was measured by turbidity meter using the technique of nephelometry, which measures the amount of light scattered at right angles to an incident light beam by particles present in a fluid sample (Sadar, 1996).

Total dissolved solids (TDS) of the sample were determined by Hanna Pocket TDS meter (HI -98130). Conductivity of the effluent sample was measured with a probe and conductivity meter. Dissolved Oxygen is the oxygen concentration in the sample which is defined as the number of milliliters of dioxygen gas (O₂) per liter of water. DO in effluent sample was measured by Winklers titration while Biological Oxygen Demand (BOD) was measured after five days BOD at 20°C was determined by standard dilution technique according to standard methods.

Chemical Oxygen Demand (COD) was measured by Potassium dichromate reflux method.

Nitrate Determination was carried out by spectrophotometer at a wavelength of 480 nm **Cataldo et al. (1975)**. Inorganic Phosphate estimation in effluent and soil was done by Vanado Molybdophosphoric Acid Method. Ammonia of the effluent was determined as Total Kjeldal Nitrogen (TKN) by **Kjeldahl method (1883)**. Chlorides of the samples were estimated by silver nitrate solution, Sodium, Potassium, Calcium and Magnesium in effluent and soil were measured by Flame Photometer (systronics). Alkalinity of the sample was measured by EDTA titration method. Salinity of the effluent was measured by salinity meter Wattson (WS-SA 287). For heavy metals i.e. Cu, Cd, Cr and Zn in effluent and soil samples were analyzed by ICP-OES spectrophotometer.

2.2 *In silico* analysis

2.2.1 Motif analysis

Motif analysis was carried out by using MEME tool and sequences with an *E*-value less than **10** the *p*-value of a hit must be less than **0.0001** are included for the analysis.

2.2.2 Protein Primary, secondary and tertiary structure prediction

Amino acid composition or linear structure of enzyme was analysed by the Bioedit standalone tool. The secondary structure was predicted by chao fassmann secondary structure prediction server (CFSSP). The tertiary structure prediction of all the enzymes was done by using RaptorX server. PMDB server was used to submit all the predicted model.

2.2.3 GRAVY (grand average of hydropathy) and molecular weight Analysis

The grand average of hydropathy value for a protein is calculated by Protein GRAVY online tool of Sequence Manipulation Suite (web server) which calculates the grand average of hydropathy value for each sequence. Increasing positive score indicates a greater hydrophobicity (**Kyte & Doolittle, 1982**).

2.2.4 Conserved domain analysis

The Conserved domain database is a large collection of protein domains. A multi FASTA format file containing protein sequences of all five enzymes was used to be searched for database at the default E-Value 1.0. The result obtained by the sequence searched was described the number of hits (domain) found in the sequences their description.

2.2.5 Phylogenetic analysis

MEGA (Molecular Evolutionary Genetics Analysis) Version 6.0.6 was used to predict the evolutionary relationship among the enzymes used in this study. Statistical parameters are considered as default and boot strap values use for tree generation was taken 1000.

3 Result and Discussion

Effluent analysis of different sites i.e. 1 to 5 is presented in the table 1. As the table shows most of the parameters were found exceeding the permissible. Annual average during the study period is shown in table 2. Through Rotated component matrix (RCM) analysis was done to explain. Results of RCM are the key output of Principal component analysis. It contains the estimates of correlation between each of the variable and estimated compartment. This helps in estimating the value of component in total value. It states that the maximum number of variables, i.e. TDS, EC, TKN and Potassium were characterized by first component, while second component is characterized by COD Nitrates and Chlorides. The third and fourth components were characterized by Turbidity and sodium respectively. The first four components (TDS, EC, TKN and Potassium) were chosen for interpretation of the data, which accounts for 71.38% of the total variance in the data set. Table 3 & 4 represents the initial eigen value, percent of variance and cumulative percent of total variance of effluent data. This shows that hydrochemical constituents of the effluent are mainly controlled by TDS and EC. It is inferred that the out of the total variance in which the first component accounts for 48.49% of the total variance. The second component explains 19.49%, the third component 7.04% and fourth component exhibit 6.295% of the total variance. Guidelines have been developed for determining the number of factors to be used. For interpreting of the data the method of Kaiser Criterion, is followed which retains only those factors having eigen values greater than 1. The first four factors having an eigen value greater than 1 has been used for interpretation. The number of eigen values can be estimated from a scree plot demonstrated in Figure 1. As shown in this figure, the eigen value sharply decrease within the first four components and then slowly stabilize for the remaining ones (Figure 3).

Table 4 represents the factor loading which were used to measure the correlation between variable and factors. Component loading higher than 0.6 is taken into consideration as suggested by **Mazlum et al. (1999)**. The components with larger variance are more desirable since they give more information about the data. The highest factor loading of TDS and EC is 0.943 and 0.925, respectively indicating the strong relationship between these two variables (**Table 4**), which is further supported by cluster analysis (Dendrogram) as shown in Figure 2.

Table 1 Total average value of various parameters of Effluent at all five sites during study period September 2009 – 2011 (mg/l)

Parameters	Site 1	Site 2	Site 3	Site 4	Site 5
Temperature	14.9 ± 3.13	14.8 ± 3.2	14.8 ± 3.2	14.9 ± 2.68	14.7 ± 3.40
pH	7.5 ± 0.43	7.33 ± 0.27	7.32 ± 0.237	7.391 ± 0.362	7.42 ± 0.731
ECmho/sec	1452 ± 32.8	1579 ± 46.28	1475 ± 77.53	1575 ± 86.54	1450 ± 134.35
TDS	1334.4 ± 90.2	1289 ± 74.34	1285.5 ± 59.58	1295.29 ± 62.64	1229.7 ± 167.04
DO	1.5 ± 0.6	1.38 ± 0.39	1.454 ± 0.4	1.53 ± 0.39	3.23 ± 2.4
BOD	286.0 ± 75.9	260 ± 30.86	264.58 ± 32.69	254.29 ± 35.64	2.11 ± 65.65
COD	713.1 ± 73.50	671.7 ± 67.4	674.25 ± 57.17	680.2 ± 62.04	521.97 ± 172.31
Na	336.6 ± 61.5	276.33 ± 52.75	301.25 ± 51.67	282.04 ± 69.81	198.91 ± 98.01
K	286.4 ± 62.8	245.95 ± 31.7	204.29 ± 71.87	229.95 ± 30.41	202.41 ± 58.86
Chloride	703.0 ± 82.2	652 ± 71.23	661.33 ± 53.6	607.75 ± 15.24	615.91 ± 79.61
Alkalinity	672.1 ± 79.5	641.58 ± 30	655.13 ± 3.59	635.41 ± 51.12	618.75 ± 109.67
Nitrate	434.0 ± 108.5	391.58 ± 88.2	360.95 ± 70.57	368.95 ± 77.76	361.95 ± 77.13
TKN	393.4 ± 61.4	338.83 ± 44.22	373.58 ± 92.47	394.65 ± 64.71	353.87 ± 131
Salinity	2.6 ± 1.4	2.71 ± 1.89	2.05 ± 0.94	2.1 ± 1.09	1.88 ± 1.08
Turbidity	67.1 ± 17.5	60.62 ± 12.3	60.04 ± 9.96	60.66 ± 12.72	59.54 ± 11.52
Phosphate	691.3 ± 55.4	584.66 ± 59.7	629.16 ± 65.1	623.83 ± 73.88	610.16 ± 63.91
Cu	0.91 ± 1.0	0.583 ± 0.63	0.6 ± 0.491	0.431 ± 0.329	0.31 ± 0.22
Zn	1.4 ± 0.4	1.9 ± 0.60	1.41 ± 0.64	1.033 ± 0.521	0.808 ± 0.575
Cd	0.2 ± 0.4	0.13 ± 0.19	0.15 ± 0.219	0.23 ± 0.27	0.3 ± 0.29
Cr	0.8 ± 0.7	1.67 ± 1.02	1.68 ± 1.09	1.35 ± 1.36	1.17 ± 1.1

Value given after ± represent Standard deviation

Table 2: Annual average of all parameters of effluent during 2009-10 and 2010-11.(mg/l)

Parameters	2009-10 Average	2010-11 Average
Temperature	13 ± 0.82	14.6 ± 0.84
pH	7.41 ± 0.16	7.45 ± 0.11
Turbidity	57.00 ± 5.20	63.98 ± 6.64
TDS	1305.78 ± 44.42	1286.92 ± 31.65
EC	591.95 ± 46.01	465.54 ± 103.91
DO	695.62 ± 981.70	1.30 ± 0.059
BOD	242.22 ± 15.39	265.00 ± 12.76
COD	684.80 ± 48.50	667.85 ± 36.05
Nitrate	402.80 ± 64.70	339.65 ± 31.46
Phosphate	610.87 ± 15.39	634.80 ± 20.20
TKN	338.35 ± 25.01	388.02 ± 13.42
Alkalinity	629.78 ± 8.89	654.77 ± 21.71
Chloride	684.90 ± 34.45	608.93 ± 29.12
Na	301.15 ± 26.78	286.23 ± 27.84
K	233.43 ± 24.48	226.62 ± 33.87
Salinity	3.07 ± 0.54	3.87 ± 0.97
Cu	0.60 ± 0.22	0.43 ± 0.12
Cr	2.22 ± 0.15	0.73 ± 0.14
Cd	0.26 ± 0.04	0.14 ± 0.01
Zn	1.44 ± 0.17	1.43 ± 0.03

Value given after ± represent Standard deviation

Table 3 Rotated Component Matrix

Parameters	Components			
	1	2	3	4
Temp	0.201	0.018	-0.117	0.757
pH	0.139	0.109	0.398	0.671
Turbidity	0.217	0.233	0.653	0.079
TDS	0.851	-0.052	0.323	0.049
EC	0.860	-0.298	0.337	0.100
DO	-0.215	-0.023	-0.759	-0.050
BOD	0.678	0.354	0.345	-0.117
COD	0.210	0.789	0.355	0.016
Nitrate	-0.314	0.841	0.087	-0.052
Phosphate	-0.716	0.522	-0.303	-0.038
TKN	0.891	-0.146	0.073	0.045
Alkalinity	0.534	0.403	0.059	0.241
Chlorides	-0.080	0.782	-0.042	0.164
Sodium	0.430	-0.045	0.596	-0.029
Potassium	0.804	0.112	0.352	-0.032
Salinity	-0.652	0.476	0.000	-0.204

Table 4 Total Variance Explained

Component (Parameters)	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% Variance	Cumulative	Total	% Variance	Cumulative	Total	% Variance	Cumulative
1	6.160	38.499	38.499	6.160	38.499	38.499	5.074	31.711	31.711
2	3.119	19.495	57.994	3.119	19.495	57.994	2.922	18.265	49.976
3	1.128	7.049	65.043	1.128	7.049	65.043	2.226	13.911	63.886
4	1.007	6.295	71.338	1.007	6.295	71.338	1.192	7.451	71.338
5	.880	5.499	76.837						
6	.682	4.260	81.096						
7	.650	4.061	85.157						
8	.551	3.442	88.599						
9	.425	2.658	91.257						
10	.367	2.293	93.550						
11	.320	1.998	95.548						
12	.219	1.366	96.913						
13	.175	1.092	98.005						
14	.167	1.044	99.049						
15	.127	.794	99.843						
16	.025	.157	100.000						

Extraction Method: Principal Component Analysis.

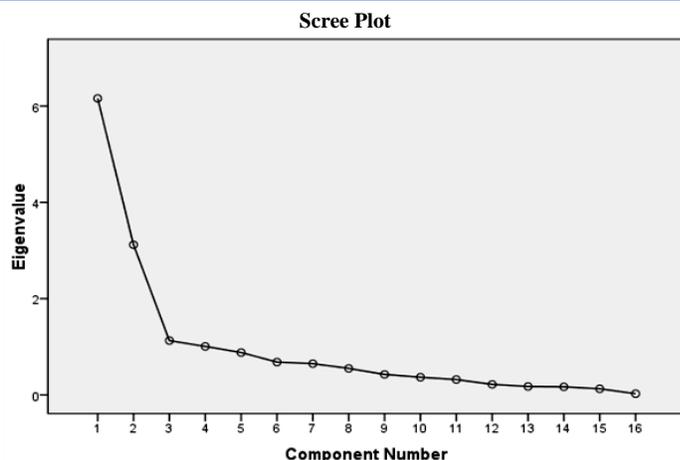


Figure 1 Scree plot for number of factor and its corresponding Eigen values of effluent

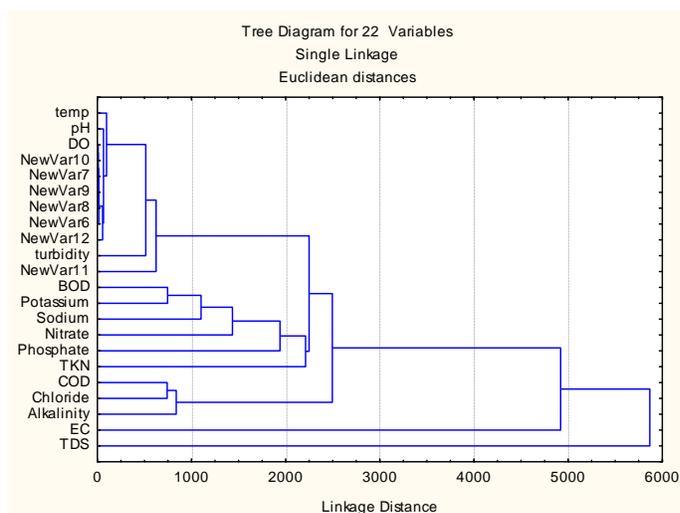


Figure 2 Cluster analysis of all parameters of effluent (Using Ward method)

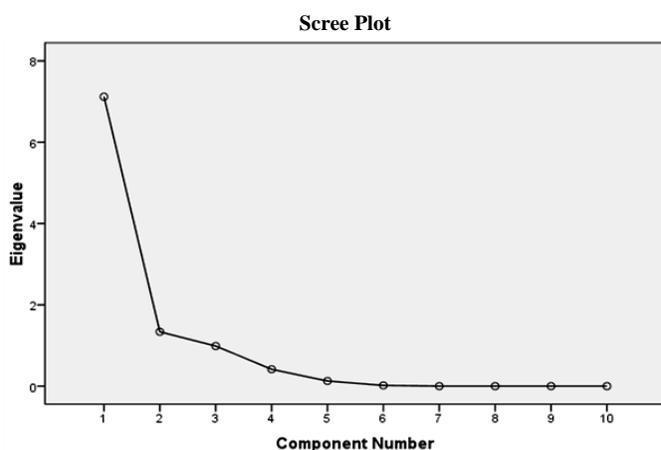


Figure 3 Screen plot for number of factor and its corresponding Eigen values of soil

Table 5 Rotated component matrix (soil)

Parameters	PCA1	PCA2	PCA3
pH	0.77	-0.53	0.14
N	0.99	0.02	0.06
PO ₄	0.97	0.02	0.22
Na	0.98	0.01	0.02
K	0.15	0.92	0.12
Cr	0.97	0.08	0.13
Cu	0.90	0.33	0.31
Cd	0.93	0.18	0.22
Zn	0.87	0.12	0.25
Pb	0.17	0.10	0.96
Eigen values	6.96	1.31	1.15
% of Variance	69.69	13.17	11.5
% of cumulative	69.69	82.86	94.41

But if the input of pollutants is higher than the purifying capacity of soil, the pollution of soil decreases microorganism effectiveness, considerably and as such nutrient on are removed from soil solution or added to soil solution, ultimately the soil always adversely affected (Table 5).

Present study puts forward the critical current situation of the different polluted sites throughout the study period. Biotechnological techniques like molecular biology, enzymes engineering, genetic engineering and bioinformatics are playing a key role in developing economically feasible and well-designed methods in production of pulp and paper sector. Nevertheless some of the studies have reported that certain enzymes are responsible for phyto-remediation of effluent and soil in these polluted sites. Yet the situation is somehow controlled by the natural cleaning process by the enzymes present in soil produced by various microbes. Considering this in mind we have selected some enzymes like ligninases, laccases, Mn peroxidases, Xylalanases and cellulases and their *in silico* analysis including domain feature, primary, secondary and structure and molecular characterization have been performed, which would be helpful for the process of detoxification of effluent in natural environment (Kirk & Chang, 1990).

3.1 In-silico analysis

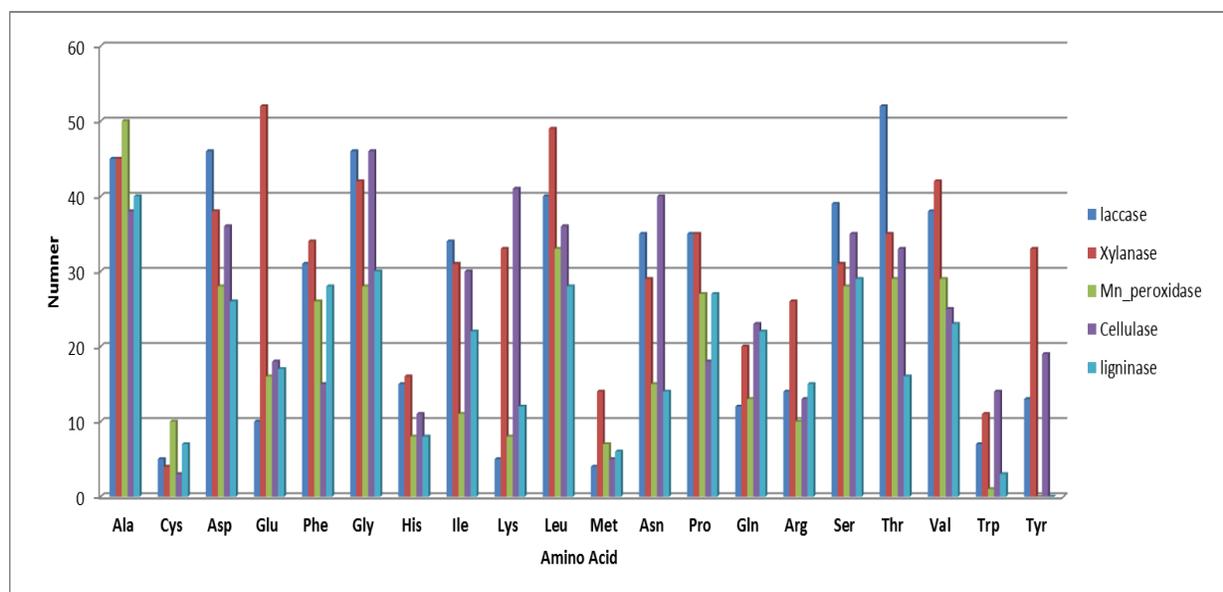
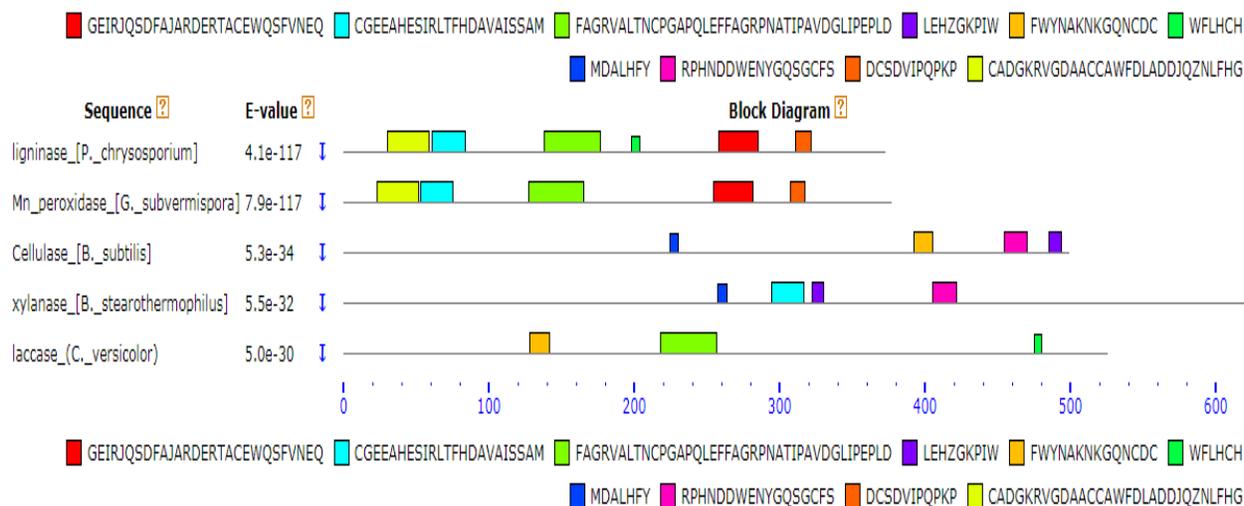
3.1.1 Motif Analysis of enzymes

In the motif study it was found that ligninase, Mn peroxidase and Laccase have similar motif structure whereas cellulase and Xylanase were found in similar motif structure. Findings of present study indicate that certain conserved residues in these motifs are involved in the waste treatment. The sequences of all motifs were indicated in different colors (Figure 4) (Bailey et al., 2009).

3.1.2 Structural Analysis

3.1.2.1 Primary Structure of Enzymes

Linear structure study of all five enzymes suggested that alanine, glycine, leucine, serine, threonine and valine are found in excess compare to other amino acids. Aspartic acid i.e. negatively charged enzyme also have a good amount in all the enzymes. Some amino acids like cystine methionine, tryptophan and tyrosine were found in less proportion (Figure 5) (Gasteiger et al., 2005; Garg et al., 2016).



3.1.2.2 Secondary Structure Prediction

Secondary structure of all these five enzymes was done by using CFSSP Server. In figure 6 it was clearly described that how much amino acids are involved in the formation of Helix, Sheet, Turn and Coil. Laccase (*C. versicolor*) which is the very first in the figure 6 and denoted by 'A' have much region of sheet (green

color). Xylanase (*B. stearrowthermophilus*), Cellulase (*B. subtilis*), ligninase (*P. chrysosporium*) and Mn peroxidase (*G. subvermispota*) are denoted by 'B', 'C', 'D' and 'E' have more region of helix formation (red color).

In the secondary structure prediction of the above enzymes (Figure 7) it was found that laccase and cellulase have high

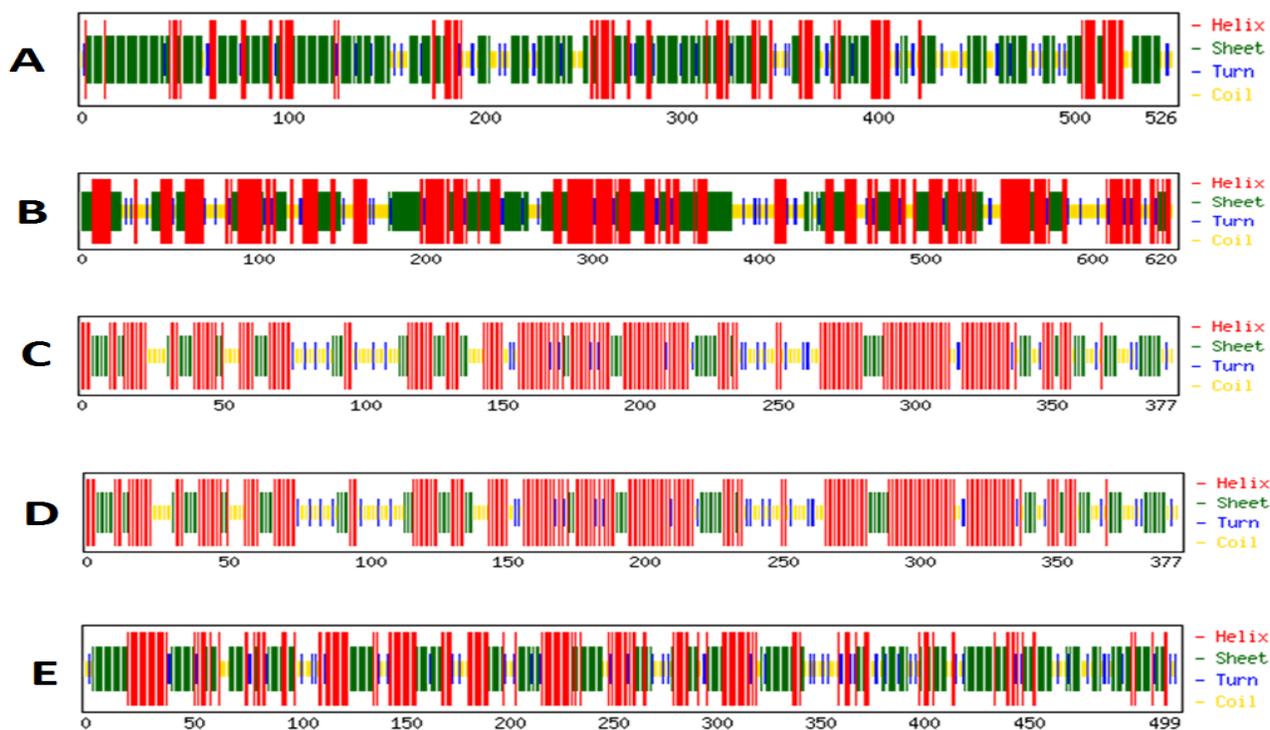


Figure 6 Secondary Structure of all five enzymes predicted by CFSSP server A, B, C, D, E. Red color shows the helix region, green shows Beeta Sheet, turn shows in blue and coil shows in yellow color.

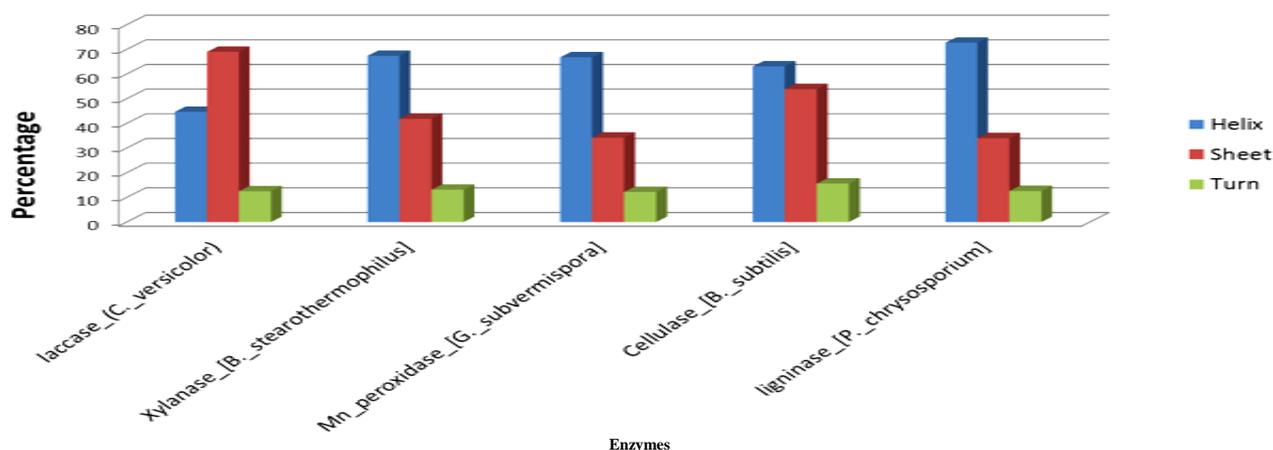
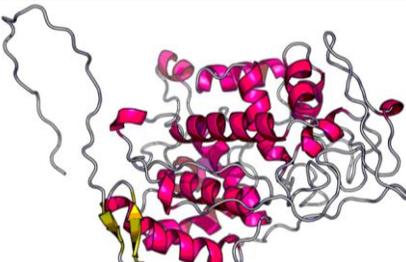
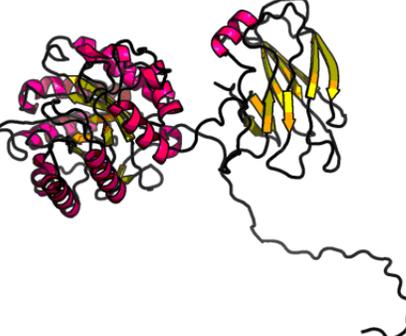
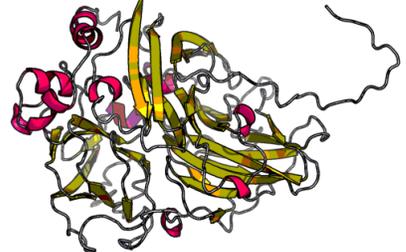


Figure 7 Helix, Sheet and Turn prortion of laccase (*C. versicolor*), Xylanase (*B. stearrowthermophilus*), Cellulase (*B. subtili*), ligninase (*P. chrysosporium*) and Mn peroxidase (*G. subvermispota*).

Table 6 Tertiary structure prediction of enzymes by RaptorX server.

S.No	Enzyme	Tertiary Structure
1	Ligninase (<i>P. chrysosporium</i>)	
2	Cellulase (<i>B. subtilis</i>)	
3	Mn peroxidase (<i>G. subvermispora</i>)	
4	Xylanase (<i>B. stearrowthermophilus</i>)	
5	Laccase (<i>C. versicolor</i>)	

region of sheet formation and others have similar secondary structure i.e. approx. equal proportion of the Helix, sheet and turn (Kumar, 2013).

Tertiary structure of all enzymes (Table 6) was predicted by the Raptor X server, which use the best template automatically with higher identity from protein data bank. The pink color in the structure is denoted the helix region and yellow color shows the sheet region in the structure (Kallberg et al., 2014).

The grand average of hydropathicity (GRAVY) for all enzymes were calculated as described by Kyte & Doolittle (1982). It was calculated by adding the hydropathy values of each amino acid residues and dividing by the number of residues in the sequence or length of the sequence. Table 7 shows that enzymes with negative GRAVY scores were significantly more abundant than those with positive GRAVY scores. The GRAVY values of most proteins were negative, which is indicating that most of them were hydrophilic. More than 6000 Dalton molecular weight, Xylanase has top position among all enzymes (Figure 8).

3.1.3 Domain Analysis

Proteins are generally composed of one or more functional regions, commonly termed domains. Different combinations of domains give rise to the diverse range of proteins found in nature. The identification of domains that occur within proteins can therefore provide insights into their function. Following conserved domains were found that may have a vital role of the waste treatment by biological ways (Table 8). Cupredoxins are involved in some biological inter-molecular electron transfer reactions. Cupredoxins are blue copper proteins that contains a mononuclear type 1 (T1) copper site. Structurally, the cupredoxin like fold consists of a beta-sandwich with 7 strands in 2 beta-sheets, which is arranged in a Greek-key beta-barrel. Multi-copper oxidases with three cupredoxin, domains are involved in Cell cycle control, cell division, chromosome partitioning, Inorganic ion transport and metabolism, Cell wall/membrane/envelope. Group of Glycosyl hydrolases are a type of enzymes, which hydrolyse the glycosidic bond between two or more carbohydrates. Heme-dependent peroxidases similar to plant peroxidases are enzymes belong to a group of peroxidases containing a heme prosthetic group, and involved in the catalysis of a multistep oxidative reactions which involved in the electron acceptor reactions. Fungal peroxidase extension region is found as an extension to a haem peroxidase domain

Table 7 GRAVY and molecular weight analysis

Enzymes	Peptide	GRAVY Score
Laccase (<i>C. versicolor</i>)	MGKFHSFVNV	-0.007
Xylanase (<i>B. stearothermophilus</i>)	MPTNLFFNAH	-0.393
Mn peroxidase (<i>G. subvermispora</i>)	MAFASLFTLV	0.122
Cellulase (<i>B. subtilis</i>)	MKRSISIFIT	-0.59
ligninase (<i>P. chrysosporium</i>)	MAFKQLFAAI	-0.073

in some fungi. Apart from the above domains, Cellulose binding domain was also found. As we know that Cellulose is an ideal matrix for large-scale affinity purification procedures. CBDs can be applied in the modification of physical and chemical properties of composite materials and the development of modified materials with improved properties (NCBI Resource Coordinators. (2013).

Phylogenetic analysis of all five enzymes was done by using MEGA 6.0 stand-alone tool and it was found that mn peroxidase and ligninase were grouped in a cluster whereas laccase, cellulase and Xylanase were fall in separate clusters (Figure 9) (Tamura et al., 2013).

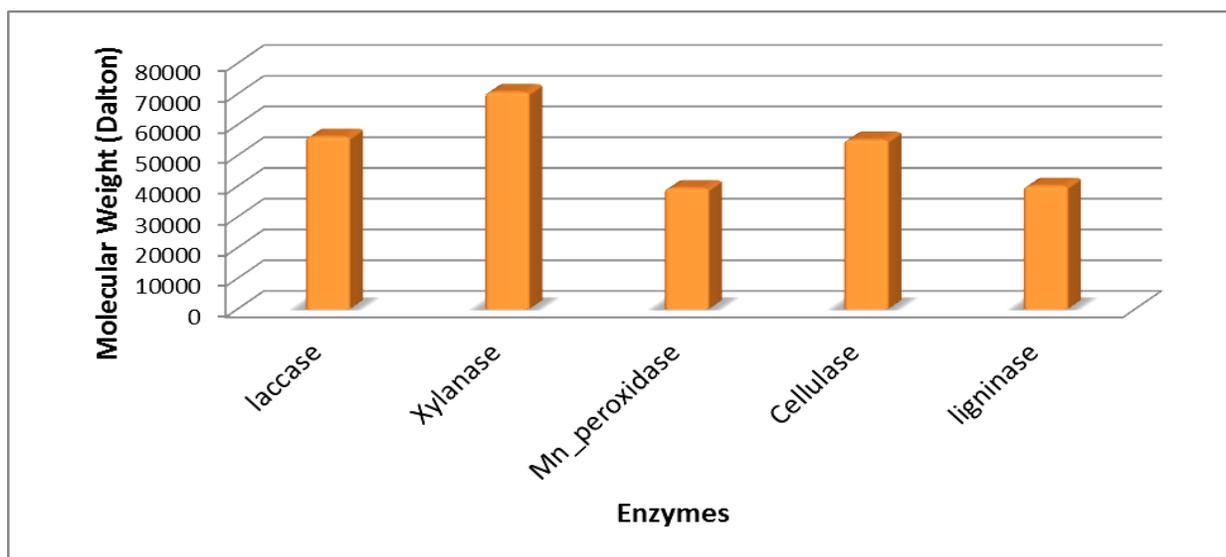


Figure 8 Pictorial representation of molecular weight of all enzymes.

Table 8 Conserved domain analysis of enzymes involved in the treatment of industrial waste.

Query	From	To	E-Value	Bitscore	Accession	Short name
Laccase (<i>C. versicolor</i>)	27	151	9.90E-81	247.251	cl19115	Cupredoxin superfamily
	348	497	1.24E-77	240.259	cl19115	Cupredoxin superfamily
	167	322	1.59E-75	235.381	cl19115	Cupredoxin superfamily
	40	496	5.78E-74	243.891	cl28276	SufI superfamily
	52	164	2.77E-10	62.3488	cl28277	PRK10965 superfamily
Xylanase (<i>B. stearothermophilus</i>)	20	437	0	639.01	cl04124	Glyco hydro 52 superfamily
	440	616	3.06E-50	174.729	cl23725	Glyco hydro 1 superfamily
Mn peroxidase (<i>G. subvermispora</i>)	22	355	1.60E-166	468.414	cl00196	Plant peroxidase like superfamily
	285	356	5.31E-41	138.602	cl12163	Peroxidase ext superfamily
Cellulase (<i>B. subtilis</i>)	50	296	1.41E-82	256.915	cl23725	Glyco hydro 1 superfamily
	9	445	3.42E-56	192.315	cl27621	BglC superfamily
	356	436	3.05E-24	95.7221	cl03026	CBM 3 superfamily
Ligninase (<i>P. chrysosporium</i>)	29	350	0	529.276	cl00196	Plant peroxidase like superfamily
	292	350	1.81E-20	83.9036	cl12163	Peroxidase ext superfamily

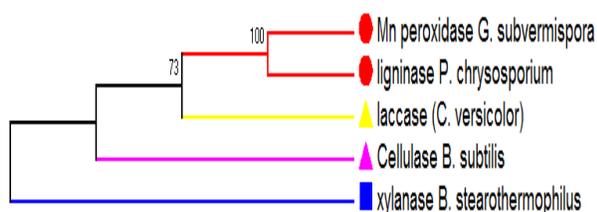


Figure 9 Phylogenetic analysis of the enzymes involved in the pollution control

Table 9 Protein tertiary structure in Protein Model database with their accessions

S.No.	PDB Structure of Enzymes	PMDB Accession
1	Laccase (<i>C. versicolor</i>)	PM0081332
2	Xylanase (<i>B. stearrowthermophilus</i>)	PM0081333
3	Mn_peroxidase_(<i>G. subvermispora</i>)	PM0081334
4	Cellulase (<i>B. subtilis</i>)	PM0081335
5	ligninase_(<i>P. chrysosporium</i>)	PM0081336

3.3 Protein model Submission

All the structures of enzymes predicted by RaptorX server were submitted to PMDB (Protein Model Database) (Table 9). Protein Model Data Base collects three dimensional protein models obtained by structure prediction methods, a joint project between CASPUR and the Biocomputing group of the Department of Biochemical Sciences of the University of Rome "La Sapienza". The PMDB Protein Model Database (Castrignano et al. 2006). These protein structures of enzymes would be further use by research community for their structure level analysis (Tiwari et al., 2016; Yadav et al., 2016)

Conclusion

All the study sites showed dismal picture of pollution status throughout the study period. Rotated component matrix and statistical analysis of effluent of all parameters showed that source of PCA1, PCA 2 is effluent and of PCA3 is earth crust. *In silico* analysis of different enzymes involved in detoxification of effluent reveals that cupredoxin family of enzymes play important role in inter-molecular transfer reaction which could break the large recalcitrant toxic molecules into simpler less toxic or nontoxic molecules. Further analysis takes us to the conclusion that the enzymes which have the large surface area, reaction sites along with the conserved binding domain (as depicted by secondary and tertiary structure and Gravy score) could be crucial in process of detoxification of paper industry effluent. *In silico* analysis has further revealed that aspartic acid as component of all

enzymes involved in detoxification. Thus aspartic acid based biochemistry may hold the key of futures research regarding detoxification and biodegradation of industrial effluents.

Conflict of Interest

Authors would hereby like to declare that there is no conflict of interests that could possibly arise.

References

- Bailey TL, Bodén M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) "MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Research* 37:W202-W208.
- Castrignano T, D'Onorio De Meo P, Ivano DC, Talamo G, Tramontano A (2006) The PMDB Protein Model Database. *Nucleic Acids Research* 34:1, D306–D309, <https://doi.org/10.1093/nar/gkj105>.
- Cataldo DA, Haroon LE, Schrader LE, Youngs (1975) VL Rapid colorimetric determination of nitrate in plant tissue by nitration of salicylic acid. *Communications in Soil Science and Plant Analysis* 6: 71–80.
- Chandra RP, Abdulsalam AK, Salim NA, Puthur JT (2010) Distribution of bio accumulated cadmium and chromium in two Vigna species and associated histological variations. *Journal of Stress Physiology and Biochemistry* 6:4-14.
- Chopra AK, Pathak C (2013) Enrichment and Translocation of Heavy Metals in Soil and Spinaceaolera Grown in Sugar Mill Effluent Irrigated Soil. *Sugar Tech, International Journal of Sugar Crops and Related Industries* 15 : 77-83.
- Garg VK, Avashthi H, Tiwari A, Jain PA, Ramkete PW, Kayastha AM, Singh VK (2016) MFPPi – Multi FASTA Prot-Param Interface. *Bioinformation* 12: 74-77.
- Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In: Walker JM (Ed): *The Proteomics Protocols Handbook*, Humana Press, Totowa, New Jersey, USA, 571-607
- Kallberg M, Margaryan G, Wang S, Ma J, Xu J (2014) Raptor X server: a resource for template-based protein structure modeling. *Methods in Molecular Biology* 1137:17-27. doi: 10.1007/978-1-4939-0366-5_2.
- Kirk TK, Chang HM (1990) Overview of Biotechnology In Biotechnology in Pulp and Paper Manufacture. Butterworth-

- Heinemann, 1-13, <https://doi.org/10.1016/B978-0-409-90192-4.50005-4>
- Kjeldahl J (1883) A New Method for the Determination of Nitrogen in Organic Matter. *Zeitschrift für Analytische Chemie* 22: 366-382. <http://dx.doi.org/10.1007/BF01338151>.
- Kumar A (2013) CFSSP: Chou and Fasman Secondary Structure Prediction server. *Wide Spectrum Research Journal* 1:15-19.
- Kyte J, Doolittle RF (1982) A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* 157:105-32.
- Mazlum N, Ozer A, Mazlum S (1999) Interpretation of Water Quality Data by Principal Components Analysis. *Turkish Journal of Engineering and Environmental Sciences* 23: 19 – 26.
- Ministry of power, GOI (2015) Normalization document and monitoring & verification guidelines (2015), Pulp and paper sector, Ministry of power, Bureau of energy efficiency, government of India.
- Mishra S, Mohanty M, Pradhan C, Patra HM, Das R, Sahoo S (2013) Physico-chemical assessment of paper mill effluent and its heavy metal remediation using aquatic macrophytes -a case study at JK Paper mill, Rayagada. *India Environment Monitoring and Assessment* 185:4347–4359 DOI 10.1007/s10661-012-2873-9.
- NCBI Resource Coordinators (2013) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research* 41: D8–D20. <http://doi.org/10.1093/nar/gks1189>.
- Qureshi MI, Muneer S, Bashir H, Ahmad J, Iqbal M (2014) Nodule physiology and proteomics of stressed legumes. *Advances in Botanical Research* 56: 1–38.
- Sadar MJ (1996) *Understanding Turbidity Science*. Technical Information Series –Booklet No. 11. Hach Company.
- Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30:2725-2729.
- Tewari PK, Batra VS, Balakrishnan M (2008) Efficient water use in industries: Cases from the Indian agro-based pulp and paper mills. *Journal of environmental management* 90: 265-273
- Tiwari A, Avasthi H, Jha R, Srivastava A, Garg VK, Ramteke PW, Kumar A (2016) Insights using the molecular model of Lipoxygenase from Finger millet *Eleusinecoracana* (L.). *Bioinformation* 12 : 156-164.
- Tornero V, Hanke G (2016) Chemical contaminants entering the marine environment from sea-based sources: A review with a focus on European seas. *Marine Pollution Bulletin* 112: 17-38. DOI: 10.1016/j.marpolbul.2016.06.091.
- Tripathi BM, Kumari P, Weber KP, Saxena AK, Arora DK, Kaushik R (2014) Influence of Long Term Irrigation with Pulp and Paper Mill Effluent on the Bacterial Community Structure and Catabolic Function in Soil. *Indian Journal Microbiology* 54: 65–73.
- Yadav N, Tiwari A, Garg VK (2016). Multiple Sequence Alignment of Model Plants Using Bioinformatics Approach. *International Journal of Agriculture Innovations and Research* 4: 2319-1473.