Summary and conclusions

SUMMARY

Lignocellulose is the most abundant bio-renewable polysaccharide on the Earth and holds enormous potential for sustainable production of bioethanol. Due to the abundant renewable source, alternative to food energy crops and huge reservoir of energy content, lignocellulose biomass has received major attention for the conversion into ethanol.
The efficient conversion of lignocellulose into ethanol requires a synergistic action of the enzymes cellulases, hemicellulases and ligninases. However, the recalcitrant nature of the lignocellulose has been the major hindrance in the lignocellulose depolymerization.

Several cellulases have been reported from a number of microorganisms, including bacteria, fungi, yeast and actinomycetes and majority of them have been obtained from the culturable microorganisms, i.e. 0.1 to 1.0 % of total microbial diversity existing in natural environments. Therefore, our current knowledge of pure-culture derived microbial enzymes for biofuel production is highly limited and search is still on for the potential cellulolytic microorganisms or enzymes for efficient bioconversion of lignocellulosic biomass. The culture-independent metagenomic approach may permit retrieval of genes encoding efficient cellulases. In this study, both culture based and metagenomic approach were used to obtain efficient cellulolytic microbes and genes from environmental samples of North-East India. The significant findings in this study are summarized below:

I. Forest litter and compost samples were collected from 20 selected sites of North-East of India and used to retrieve culturable cellulose degrading microbes and genes encoding cellulases using culture dependent and independent approaches, respectively.

II. Cellulolytic microorganisms including bacteria, fungi and actinobacteria were isolated from forest litter and compost samples on carboxymethyl cellulose (CMC) containing agar media (M-9 minimal medium containing 0.2% CMC). A total of 207 cellulose degrading microbes were obtained and tested for cellulase
activity on CMC agar plates using Gram’s iodine staining of which 101 isolates had cellulase activity (74 bacteria, 18 fungi and 9 actinobacteria).

III. Best five bacterial isolates MNP-60, DSNP-28412, SPLNP-26512, GMTS-29 and NMNP-12512 were selected and identified as *Pseudomonas resinovorans* strain MNP-60, *Paenibacillus taichungensis* strain DSNP-28412, *Pseudomonas aeruginosa* strain SPLNP-26512, *Arthrobacter phenanthrenivorans* strain GMTS-29 and *Pseudomonas mendocina* strain NMNP-12512.

IV. The selected bacterial strains (MNP-60, DSNP-28412, SPLNP-26512, GMTS-29 and NMNP-12512) were tested for their filter paper degradation abilities. All the bacterial strains were able to degrade filter paper of which *Pseudomonas resinovorans* strain MNP-60 and *Paenibacillus taichungensis* strain DSNP-28412 degraded filter paper faster compared to the reference strain *Cellulomonas cellulans* MTCC 23.

V. Solid state fermentation of rice straw with bacteria exhibited the production of major cellulases. The bacterium, *P. resinovorans* strain MNP-60 exhibited lignin peroxidase (11.40 IU/g rice straw) and Mn peroxidase (21.61 IU/g rice straw) activities and higher total cellulase (54.39 IU/g rice straw), endoglucanase (185.40 IU/g rice straw), and β-glucosidase (102.47 IU/g rice straw) activities compared to the reference strain MTCC 23.

VI. FTIR spectroscopy based study of bacteria treated rice straw indicated the action of bacteria on lignocellulose components of rice straw. The FTIR spectrum of rice straw changed after the bacterial treatment. The bacteria *P. resinovorans* strain MNP-60 and *P. taichungensis* strain DSNP-28412 treatment caused most structural changes in the rice straw.

VIII. The selected fungal strains (MNP-C, ONP-13512, TWRF-10, MBL-115, BLRF-115, RWS-30, TRF-27 and BLPNP-17) were used for solid state fermentation using rice straw and lignolytic and cellulolytic activities were determined. The fungi, *P. simplicissimum* strain TRF-27 had highest laccase (248.53 IU/g rice straw) and Mn peroxidase (168.5 IU/g rice straw) activities and fungi, *F. equiseti* strain TWRF-10 had higher endoglucanase, exoglucanase and endoxylanase activities compared to the reference cellulolytic fungus *F. solani* strain MTCC 10158.

IX. The SDS-PAGE based fungal proteins analysis indicated different secretory protein profiles of the fungal strains. The secretory proteins of the strain *F. solani* MTCC 10158 were different from the *F. solani* ONP-13512, but similar to the *P. citrinum* strain BLPNP-17. On the otherhand, the closely related strains MNP-C and MBL-115 belonging to *F. equiseti* had similar secretory proteins.

X. Scanning electron microscopy (SEM) analysis of fungi treated rice straw revealed the fungal colonization and structural changes in rice straw. Characteristic mycelial growth and attachment of fungal mycelia on rice straw caused changes in the surface morphology with exposed cellulose microfibrils in fungi treated rice straw.
XI. X-ray diffraction (XRD) analysis of fungi treated rice straw revealed the reduction in the crystallinity index (CrI) of fungal treated rice straw by 19-65% indicating degradation of crystalline cellulose as well as amorphous components of the cellulose, including lignin and hemicellulose. Fungi Penicillium simplicissimum strain TRF-27 caused highest reduction of CrI (65%), followed by P. citrinum strain RWS-30 (62%).

XII. The FTIR analysis of fungi treated rice straw indicated the changes in the functional groups of lignocellulose components of rice straw. The strains of F. equiseti (MNP-C, MBL-115 and TWRF-10) had similar effect on rice straw degradation and the strains of Penicillium (MNP-C, MBL-115, TWRF-10) and Mucor (BLRF-115) had similar effects on rice straw irrespective of their species suggesting that though, these strains have originated from geographically isolated places had similar action on rice straw.

XIII. A principal component analysis (PCA) of FTIR data (% absorbance) of fungi treated rice straw indicated the grouping of different strains of Penicillium sp. and F. equiseti. This revealed that the characteristic action of the fungal strains on rice straw is linked to their genetic relatedness.

XIV. Metagenomic approach was used to retrieve the genes encoding cellulases from unculturable microorganisms present in the forest litter and compost environment. An improved DNA extraction protocol for extracting high molecular weight DNA (40 kb in size) with less humic acid contamination was optimized.

XV. Four fosmid libraries using metagenomic DNA extracted from compost samples were constructed and a total of approximately 4.9×10^7 clones with 1.9 Tb of metagenomic DNA were obtained.
XVI. Function based approach was adopted for the screening of cellulase positive fosmid clones. A total of 8,900 transformants were screened on cellulose containing agar plates and six fosmid clones (pFOS-C1, pFOS-C2, pFOS-C3, pFOS-C4, pFOS-C5 and pFOS-C6) exhibited clear halo zones of hydrolysis on agar plates. These six fosmid clones further tested for qualitative filter paper degradation to identify the clones with multiple cellulolytic enzymes. Among six clones, only fosmid clone pFOS-C1 degraded the filter paper after 20 days of incubation. The clone, pFOS-C1 was selected for sequencing of the metagenomic DNA insert.

XVII. The metagenomic DNA insert from pFOS-C1 was sequenced and after removal of vector contamination, DNA insert size was confirmed as 34.9 kb with G+C content of 70.2%. The open reading frame (ORF) finder analysis revealed a total number of 198 ORFs (genes) in the DNA insert of which 3 ORFs, ORF67, 21 and 47 had weak homologies with glycosyl hydrolase family proteins i.e. 31% homology with the glycosyl hydrolase family of Solirubrobacter sp. URDH0082, 29% homology with the glycoside hydrolase family of Zobellia galactanivorans and 39% homology with the endoglucanase of Sphingobacteriales bacterium 41-5, respectively.

XVIII. ProtParam (ExPASy) analysis indicated the acidic nature of the proteins with pI 4.43 (ORF67), 5.0 (ORF47) and 5.81 (ORF21). The promoter prediction in ORFs encoding cellulases using BPROM tool showed the presence of -35 and -10 box at the upstream of the ORFs with start and stop codons in the ORFs. The ORF67, ORF21 and ORF47 consists of -35, -10 box as -35(TTCCCG) -
Summary and conclusions

10(GCCGAACAT), -35(TTCAAA) -10(ATGCATCAA) and -35(TTGTCA) - 10(CGGGACAAT) respectively.

XIX. One of the ORF (ORF67/ghc1) of size 1128 bp encoding 375 amino acids was successfully cloned in pET28a(+) vector and expressed in E. coli BL21(DE3). The protein, GHC1 expressed in the extracellular fraction. The partially purified GHC1 protein with expected size of 40 kDa had carboxymethyl cellulase (CMCase) activity at optimal pH 5.0 and 40°C. The specific CMCase activity (45.07 IU/mg) of the GHC1 was found to be 17.61 and 7.88 times higher than the commercial cellulases (Cellulase from Aspergillus niger, C1184, SIGMA) (2.55 IU/mg) and the reference bacterium, C. cellulans strain MTCC23 (5.72 IU/mg), respectively.

XX. Multiple sequence alignment of amino acids of GHC1 with known carboxymethyl cellulases retrieved from NCBI showed poor sequence similarities (4.2%, 16/376 amino acids). The phylogenetic tree of GHC1 protein with the hypothetical and GH family proteins showed outgrouping of the GHC1 indicating that GHC1 is a new or novel member of endo-type cellulase family protein.

XXI. The homology model of GHC1 protein predicted by Phyre2-protein model recognition server had 52% of β sheets (195 of 375 amino acids) and 1% of α helix (3 of 375 amino acids). The PROCHECK and Ramachandran plot based stereochemical qualities validated the quality of the homology model.

XXII. Molecular dynamic simulations for GHC1 were performed using GROMACS 4.6 and the internal energy of GHC1 was minimized to -7.87×10^5 KJ/mol using steepest descent algorithm indicating the stabilized protein. Docking of energy minimized GHC1 with ligand carboxymethyl cellulose (CMC) resulted the
interaction of CMC in the catalytic pocket of the protein. The residues of GHC1, Ser$^{162}$, Ser$^{229}$, Ala$^{228}$, Val$^{230}$ and Gly$^{289}$ were involved in the interaction by forming hydrogen bonds with Ala$^{228}$ and closest contacts with Ser$^{162}$, Ser$^{229}$, Val$^{230}$ and Gly$^{289}$, respectively.
CONCLUSIONS

Five bacteria and eight fungi isolated from composts and forest litters of North-East India were evaluated for lignocellulolytic activities. The bacteria, *P. resinovorans* strain MNP-60 and *P. taichungensis* strain DSNP-28412 had ability to act on untreated rice straw without delignification pretreatment as revealed by FTIR analysis. The fungi, *F. equiseti* strain TWRF-10 and *P. simplicissimum* strain TRF-27 had high cellulase and ligninase activities, respectively. The overall chemical changes caused in rice straw due to the fungal treatment could not be explained by the conventional enzyme assays as revealed by XRD and FTIR analysis. These strains should be further tested for industrial use. In culture independent approach, a novel carboxymethyl cellulase gene *ghc1* was isolated from the metagenome of a compost sample collected from Tawang reserve forest, Arunachal Pradesh, India and its enzymatic activities were characterized. The recombinant GHC1 had 17 fold higher activity than the reference commercial cellulase indicating its potential for industrial applications.