

Full Length Research Paper

Cloning and characterization of cDNA encoding xyloglucan endotransglucosylase in *Pennisetum glaucum* L.

Manoj Kumar Srivastava¹, Chandrabhan Yadav², Vishnu Bhat² and Suresh Kumar^{1,3*}

¹Division of Crop Improvement, Indian Grassland and Fodder Research Institute, Jhansi - 284 003, India.

²Department of Botany, University of Delhi, Delhi-110007, India.

³Department of Horticulture and Landscape Architecture, Purdue University, West Lafayette, Indiana - 47907 2010, USA.

Accepted 1 July, 2011

Biomass production in plant is directly related to the amount of intercepted solar radiation by the canopy and available water to the plant. Growth and development of leaves, especially under drought condition, is therefore major determinant of crop productivity. Xyloglucan endotransglucosylase (XET) plays important role in growth and development of plants. XETs are a family of enzymes that mediate construction and restructuring of xyloglucan cross-links, thereby controlling the mechanical properties of cell wall. We cloned complete cDNA of an XET from pearl millet (*Pennisetum glaucum* L.) and characterized it using *in silico* comparative genomics and activity assays. The cloned cDNA was 1266 bp in length, encoding a protein with 291 amino acids having signal peptide targeting it to the cell wall. The protein showed xyloglucan endotransglucosylase activity but no hydrolytic activity, therefore, named as *PgXET1* as per the convention. The comparative genomics revealed that the functional sites of the enzyme (XET) were highly conserved. Evolutionary studies using phylogenetic tree indicated its grouping with XETs from maize (with >95% bootstrap support), barley, rice, etc. This is the first report on cloning and characterization of an XET (*PgXET1*) from pearl millet, an important dual-purpose crop.

Key words: Xyloglucan endotransglucosylase, *Pennisetum glaucum*, pearl millet, primary cell wall, cell expansion, drought tolerance.

INTRODUCTION

Plant cells are enclosed by cell walls that define shape and size of cells and mediate cell-to-cell contact. Cell wall is an important structural component that glues cells together, provides mechanical strength and rigidity and

serves as a barrier against pathogens (Cosgrove, 2000). The dynamics of plant growth, morphogenesis and differentiation requires concomitant modifications of cell wall. Cell shape is highly dependent on cell wall composition. Primary cell wall consists of three co-extensive polymer networks: the cellulose-xyloglucan framework, pectin and structural proteins (McQueen-Mason, 1997). In the growing plant, it is composed of complex carbohydrates (~90%) and proteins. Xyloglucan is an important constituent and plays a crucial role in cell enlargement during plant growth. Xyloglucans in monocotyledonous plants contain no fucose, less xylose and much lesser galactose than those found in dicotyledons. The major proteins found in the primary cell wall are hydroxyproline rich glycoproteins (extensins),

*Corresponding author. E-mail: sureshkumar3_in@yahoo.co.uk or kumar112@purdue.edu Tel: +91 510 2730908. Fax: +91 510 2730833.

Abbreviations: XET, Xyloglucan endotransglucosylase; RT-PCR, reverse transcription polymerase chain reaction; cDNA, complementary deoxyribonucleic acid; RACE, rapid amplification of cDNA ends; NCBI, national center for biotechnology information; ORF, open reading frame.

arabinogalactan proteins, glycine-rich proteins, proline-rich proteins and some other chimeric proteins (Carpita et al., 1996).

The primary cell walls of grasses have different chemical composition when compared with other flowering plants. They vary in complex glycans, nature of gel matrix and the types of aromatic substances and structural proteins. A mixed link 1,3:1,4- β -glucan is the major hemicellulose found in the cell wall of grasses. This mixed link glucan might inhibit wall cross linking and act as filler. However, in spite of these structural differences, the developmental physiology and their responses to endogenous and exogenous growth regulators are similar to those in other flowering plants. Thus, cell wall in grasses performs the same wall functions as in other plants but with a different chemical composition and architecture (Carpita et al., 1996). The cell wall contains numerous enzymes which can hydrolyze various components of wall matrix. This activity is important for wall extension, wall loosening, in defense mechanism, signaling and polysaccharide breakdown to provide energy. These cell wall modifying enzymes regulate structural changes in the polymer network resulting in the plasticity of cell wall (Xu et al., 1996). A continuous enzymatic modification in the cell wall is required during plant growth and development, environmental stresses and pathogenic infection. Therefore, biosynthesis and differentiation of cell walls are precisely regulated in a temporal, spatial and developmental manner. Many of the growth related studies have focused on breakdown of matrix glucans, namely xyloglucans and 1,3:1,4- β -glucan, found in the cell wall of seedlings (Carpita et al., 1996).

A class of enzymes known as xyloglucan endotransglucosylases (XETs) has the potential to enzymatically modify wall components. Although, their biochemical activity has been defined (Campbell and Braam, 1999), the physiological roles remain undefined. XETs are generally encoded by large gene families. To understand their physiological roles, the diverse regulation of the genes and properties of the proteins are being determined (Xu et al., 1996). XET cleaves the xyloglucan polymers at an internal site and joins the newly generated reducing ends to other xyloglucan chains. It catalyses transglycosylation of xyloglucan which has been reported to mediate the cross linking of cellulose microfibrils in cell wall and thereby controls the cell wall relaxation (Cosgrove, 2005). XET activity has been directly related to growth of the plants and elongation of the organs. Thus, XET genes are important for bioengineering of plant growth and cell wall.

Expression of XET genes is differentially regulated by several developmental and environmental factors. It is commonly observed that root/shoot ratio in plants increase when water availability is limited. The physiological and molecular mechanisms that assist root growth under drought conditions have been reviewed by Wu and

Cosgrove (2000). XETs, in association with other wall enzymes, play important role in adaptation of plants to drought conditions (Wu et al., 1996). Maize seedlings adapt to low water potential by making the walls in the apical part of the root more extendible. This is accomplished, in part, by an increase in expansion activity and in part by other more complex changes in the wall (Wu et al., 2001).

XET proteins are of moderate size (~33 kD) and N-glycosylated. XETs contain a catalytic domain (DEID-I/F-EFL) which is homologous to the catalytic domain of several bacterial endo-1,3:1,4- β -glucanases. Point mutation leading to substitution of the first E by Q resulted in an inactive protein (Campbell and Braam, 1998). Different XETs in plant may vary in their specificity for transglycosylation and/or hydrolysis reactions. This variation has been attributed to the structural differences in and around the catalytic site (Nishitani and Tominaga, 1992; Fanutti et al., 1993).

In most crop species, dry matter production is directly related to the amount of intercepted solar radiation by the canopy. Growth and development of leaves is therefore a major determinant of crop productivity. This is of specific significance for forage crops, where the greatest proportion of the shoot is harvested and used as feed for ruminants. Leaves of forage plants are therefore not only the photosynthetic active tissue, but also the harvestable product. In managed grassland systems forage plants are regularly subjected to a more or less severe defoliation by grazing or mowing, resulting in a complete or partial loss of the leaf area. To avoid prolonged periods with a reduced interception of the solar radiation, a rapid leaf growth is a prerequisite (Reidy, 2001). Factors associated with the growth of leaves are thus, keys to improve productivity as well as quality of the forages produced.

Pearl millet [*Pennisetum glaucum* (L.) R.Br.] is a staple food and forage (dual-purpose) crop in semi-arid regions of Africa and the Indian sub-continent. In the world, importance wise pearl millet occupies sixth rank after wheat, rice, corn, barley and sorghum (Rai et al., 1999). In India, it is widely grown in Rajasthan, Gujarat, Maharashtra, Haryana, Andhra Pradesh, Uttar Pradesh and Punjab (www.milletindia.org). Mainly green fodder of pearl millet is fed to the cattle and the green fodder yield varies from 35 to 40 t/ha (Gangaiah, 2008). Though it yields 6 to 7 t/ha of dry fodder (*karvi*), it is least preferred by animals. Being an important forage crop, pearl millet improvement programmes need to concentrate on increasing biomass production per unit area with decreasing availability of water for irrigation, particularly in forage crops. Since XET activity is directly related to the growth of plant and elongation of the organs, XET genes are important candidate which may be deployed for engineering forage crops to improve plant growth and cell wall components (Burstin, 2000). Due to the specific significance of leaf growth and drought tolerance for

forage crops, we aim at studying the XET genes in pearl millet in regulation of these traits. We report, cloning and characterization of cDNA for an XET enzyme in pearl millet (*P. glaucum*) and studying its evolutionary relationship with known XETs from other species (both monocots and dicots).

MATERIALS AND METHODS

A month old plants of pearl millet (*P. glaucum*, cv. JHPM 05-02) from IGFR Research Farm, Jhansi, were used as the source of plant material. Young leaf tissues (~100 mg) from the plant were used for RNA isolation. Leaf tissues were ground into fine powder in liquid nitrogen using a pestle and a mortar. Total RNA was isolated using RNeasy Plant Minikit (QIAGEN GmbH, Germany) as per the manufacturer instructions. RNA was quantified using UV visible spectrophotometer (NanoDrop 2000, Thermo Scientific, USA).

Cloning of a *P. glaucum* cDNA

About 10 µg of total RNA was used for reverse transcription PCR using M-MuLV Superscript reverse transcriptase (Fermentas GmbH, Germany). The RT-PCR reaction was set up according to the manufacturer's instructions. The resulting cDNA served as template in PCR reaction using degenerate primers set (designed from the conserved nucleotide sequence of accession numbers X91659, X93174, X93175, FJ940680 and 606356). Forward primers 5' CACGAYGAGATCGACTTCGAGTT 3' and reverse primer 5'GAGGCCARTCGTCGGCGTTCCA 3' were used to amplify a potential XET cDNA encoding part of the enzyme containing active site. The PCR conditions were: 94°C for 5 min followed by 39 cycles of amplification (94°C for 1 min, 56°C for 1 min and 72°C for 1 min) and a final extension at 72°C for 10 min. The PCR product was checked on 1.6% agarose gel and a band of interest (~300 bp) was eluted from the gel using GenElute Gel Extraction Kit (Sigma-Aldrich). The gel purified PCR product was cloned into pGEM-T Easy vector (Promega Corporation, WI, USA). The recombinant plasmid was used for transformation of competent cells of DH5 strain of *Escherichia coli*. The transformed cells were screened by blue-white screening method using X-gal and IPTG. Ten white colonies were randomly selected and the plasmids were isolated from the bacterial culture using QIAprep Spin Miniprep (QIAGEN GmbH, Germany) plasmid isolation kit. The plasmids from the recombinant clones were linearized by restriction digestion with *Pst*I, while *Eco*RI was used to release the insert from recombinant plasmid. Once confirmed about the presence of the insert in the recombinant plasmid, two randomly selected clones were sequenced by ABI Prism-373 DNA sequencer using the flanking universal primers SP6 and T7. The nucleotide sequence data received from the DNA sequencer was processed to remove vector sequences. Consensus sequence of the fragment was determined by comparing both strand sequences for the duplicate samples. The coded amino acid sequence was deduced and the presence of the proposed catalytic site in angiosperm XET, that is, DEIDFELG or a variant, was confirmed (Campbell and Braam, 1999).

The 3'-end of the cDNA was obtained by 3'-RACE using a gene-specific primer and oligo-dT₂₆. The 5'-end was cloned using the 5'-RACE kit (Invitrogen, CA, USA) with gene-specific primer. The partial cDNAs were combined to make composite cDNA. To confirm the nucleotide sequence, full-length cDNA was amplified with primers designed from both ends of the composite cDNAs and total

RNAs. In general, the protocol of Harada et al. (2010) was used with necessary optimization followed. The processed cDNA (complete cds) sequence was submitted to the NCBI GenBank (<http://www.ncbi.nlm.nih.gov/genbank>) with the accession number HQ416697.

In silico sequence analysis

The amino acids sequence encoded by the cloned cDNA was deduced using ExPASy-Translate tool from ExPASy Proteomics Server of Swiss Institute of Bioinformatics (www.expasy.ch/tools/dna.html). Sequence similarity searches were performed using BLAST at the National Center for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and similar sequences were downloaded for comparative *in silico* analyses. Multiple sequence alignments for DNA and protein were performed using ClustalX (Thompson et al., 1997) and manually visualized using BioEdit graphical view (Hall, 1999). Signal peptide and cleavage site prediction was done using neural networks (NN) and hidden Markov models (HMM) trained on eukaryotes at SignalP 3.0 server of Technical University of Denmark (<http://www.cbs.dtu.dk>). Prediction of sub-cellular location of the protein was confirmed using Plant-mPLoc (Chou and Shen, 2010; <http://www.csbio.sjtu.edu.cn/bioinf/plant>).

Expression and enzyme activity assay

The *PgXET1* cDNA was amplified using Phusion[®] High-Fidelity DNA polymerase (Finnzymes, Finland) without its signal peptide, in-frame with the factor secretion signal peptide of the pPICZ A vector (Invitrogen, CA, USA) and with the stop codon. The PCR product was subcloned in pPICZ vector in reading frame A. *Pichia pastoris* (strain GS115) was transformed with *PgXET1* expression vector by electroporation as per the manufacturer's instructions (EasySelect *Pichia* Expression Kit, Invitrogen). Colony PCR was used to confirm the expression construct in the yeast genome. A PCR positive clone was used to grow starter culture in BMGY medium with zeocin (100 mg/l) at 30°C overnight. Large culture was inoculated with 0.1% starter culture and grown overnight. The cells were harvested by centrifugation at 5000 rpm for 5 min and resuspended in one-fourth (large-culture) volume of buffered methanol-complex medium (1% methanol). After 3 days induction, the yeast cells were removed by centrifugation and supernatant was used to purify the protein by (NH₄)₂SO₄ precipitation followed by dialysis against 2 changes of 12 h each of buffer A (50 mM sodium acetate buffer pH 5, 50 mM NaCl, 10% glycerol) at 4°C. The purified protein was run on SDS-PAGE to check purity and size of the protein and then used for XET activity by dot-spot activity assay (Van Sandt et al., 2006).

The purified *PgXET1* was assayed for xyloglucan endo-transglucosylase activity using the procedure described by Fry (1997) with necessary optimization. 10 µl of the purified *PgXET1* was spotted on xyloglucan-containing test paper along with the sulphorhodamine-labelled xyloglucan oligosaccharide (XGO-SR = acceptor substrate). As a control, trisaccharide-SR (non-XET substrate) was used in a parallel experiment. The spotted test-paper was sealed in acetate envelope to avoid drying and incubated at 28°C overnight. The paper was then washed in equal (v/v) mixture of 90% formic acid, ethanol and water for 4 h followed by a 5 min wash in water so as to remove non-reacted fluorescent xyloglucan oligosaccharides. The test paper was observed under UV light and photographed.

The *PgXET1* was also assayed for hydrolase activity using the iodine staining method as described by Suda et al. (2003). About

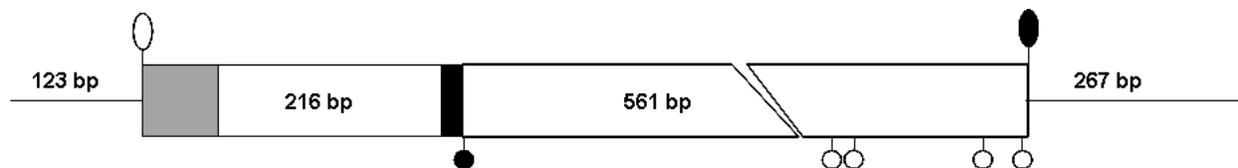


Figure 1. Structure of the XET cDNA (based on sequence analysis) cloned from leaves of *P. glaucum*. Open ellipsoid represents ATG start codon, light shaded rectangle – signal peptide, dark shaded rectangle – active site motif, filled circle – N-glycosylation site, open circles – cysteine residues, closed ellipsoid– TGA stop codon.

50 ng recombinant *PgXET1* protein was used in the assay. To measure the contribution of substrate hydrolysis, control reaction without xyloglucan heptasaccharide was included. The reaction mixture was incubated in the dark for 30 min at room temperature to allow colour development and then absorbance was taken at 630 nm.

Phylogenetic relationship

The phylogenetic relationship of *PgXET1* with other known XETs was studied by generating a phylogenetic tree using full length protein sequence of XETs from 25 different plant species. The GenBank accession numbers of 30 different XETs are: *Arabidopsis thaliana* (D16454), *Glycin max* (D16455), *Hordeum vulgare1* (X91659), *H. vulgare2* (X93175), *H. vulgare3* (CAA63662), *Triticum aestivum* (D16457), *Solanum esculentum1* (CAA58003), *Gossypium hirsutum* (D88413), *Zea mays1* (U15781), *Z. mays2* (ADB51615), *Capsicum annum1* (ABD96607), *C. annum2* (ABD96608), *C. annum3* (ABD96609), *Striga asiatica* (ABD98046), *Oryza sativa* (BAD53910), *Asparagus officinalis1* (AF223420) *A. officinalis2* (AF224319), *Musa accuminata* (ABL10090), *Vigna angularis* (D16458), *Nicotiana tabacum* (ADB41673), *Solanum tuberosum* (CAJ77496), *Fagus sylvatica* (CAA10231), *Litchi chinensis* (ABK30789), *Cucumis sativus* (CAD87535), *Festuca pratensis* (CAC40807), *Brassica rapa* (ACB86852), *Dahlia pinnata* (ADJ67812), *Pisum sativum* (BAB17788), *Rosa chinensis* (ADB55706) and *Medicago truncatula* (ABI37007). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). Bootstrap analysis was performed using 5000 replicates. An optimal tree with the sum of branch length = 4.08479105 was obtained. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (5000 replicates) were figured out next to the branches (Felsenstein, 1985). The evolutionary distances were computed using the p-distance method (Nei and Kumar, 2000) and presented in the units of the number of amino acid differences per site. The protein alignment was analyzed using MEGA 5 (Tamura et al., 2011) to obtain a bootstrapped neighbor-joining tree based on synonymous or non-synonymous changes. The tree was drawn to scale, with the same units for branch lengths and evolutionary distances. All ambiguous positions were removed for each sequence pair.

RESULTS

Cloning of cDNA for *PgXET1*

cDNA encoding a putative XET was cloned from leaf tissues of *P. glaucum* combining RT-PCR, 5'-RACE and 3'-RACE techniques. Use of the degenerate primers, designed from the conserved sequence of the active site of the enzyme, resulted into amplification of a ~300 bp

fragment. The flanking 5' and 3' sequences were successfully cloned using the gene-specific primers and RACE techniques. After getting the composite cDNA sequence information, the complete nucleotide sequence for the ORF of the gene was cloned by amplification of full-length cDNA using PCR. The cloned XET cDNA from *P. glaucum* was named as *PgXET1* (according to the 9th Cell Wall Meeting held in Toulouse, France, 2001; Yokoyama and Nishittani, 2001; Rose et al., 2002). The cDNA (complete cds) sequence submitted to the NCBI GenBank was assigned with the accession number: HQ416697.

In silico structural analysis of *PgXET1*

Predicted structure of the *PgXET1* cDNA is graphically represented in Figure 1, showing a 5'-flanking sequence of 123 bp, open reading frame (ORF) and a 3'-flanking sequence. The complete cDNA was found to be 1266 bp long and the predicted protein consisted of 291 amino acids, corresponding to the calculated molecular weight of 32.7 kDa. The predicted protein was found to have a potential signal peptide sequence from amino acids 1 to 24 (Figure 2). The most likely cleavage site was predicted between amino acids 24 and 25 ($A^{22}Q^{23}G^{24} - G^{25}N^{26}F^{27}$). Among the 12 sub-cellular locations of plant proteins (Chou and Shen, 2010), the predicted location of the *PgXET1* protein was found to be the plant cell-wall.

An active site motif (DEIDFEFLG) was found to be located between amino acids 97 to 105 followed by N-X-T/S, a potential site for N-glycosylation (amino acids 106 to 108). Four cysteine residues (amino acids 220, 229, 274 and 288) were found to be well conserved among all the species. *PgXET1* contains TGA stop codon followed by several alternative stop codons. Arginine, glutamic acid and aspartic acid are found conserved at well defined positions. The amino acids E^{110} , R^{131} involved in N-glycan interactions are conserved or substituted to homologous amino acids E^{110} , H^{115} , R^{131} , F^{140} , Y^{183} , D^{191} , W^{192} and G^{196} , were also found to be conserved in *PgXET1*.

Alignment search (using BLASTp) of amino acid sequence of the *PgXET1* with the available protein database showed its maximum homology with *Zea mays* XTH protein (Xyloglucan endotransglucosylase/hydrolase protein 23; accession No. NP001151103), followed by

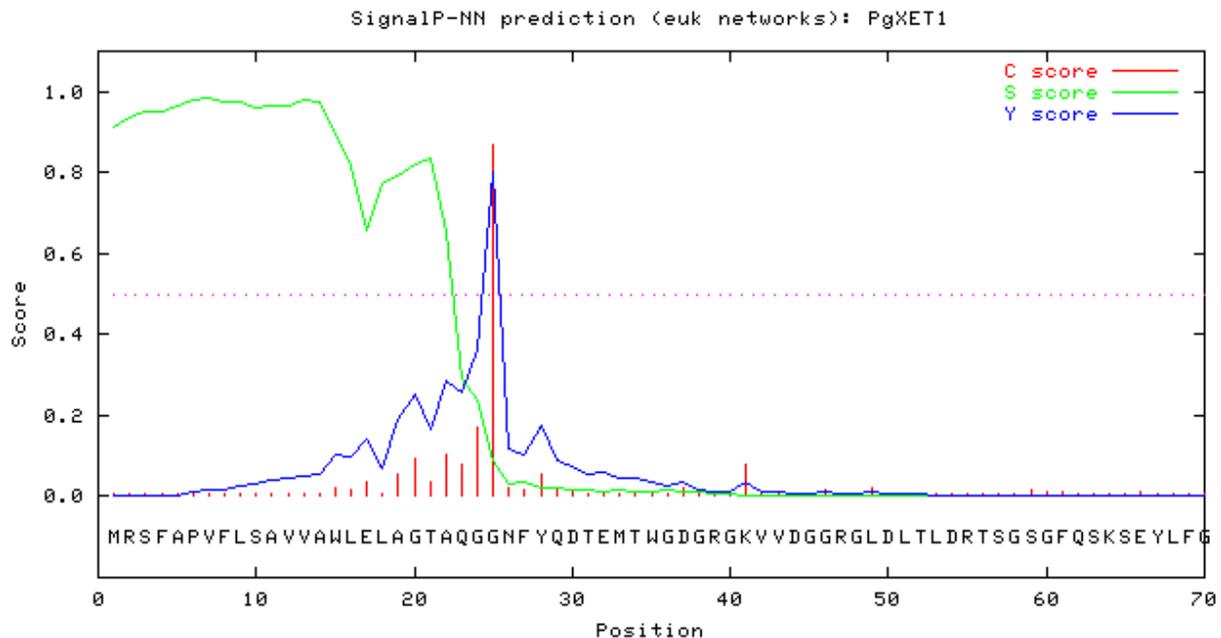


Figure 2. Prediction of signal peptide and its cleavage site using neural networks (NN) model on SignalP 3.0 server (<http://www.cbs.dtu.dk>). The most likely site of cleavage is between amino acids G²⁴ and G²⁵ (Chou and Shen, 2010).

XET sequences from *Zea mays* (*Z. mays2*). Comparative analysis of the ZmXTH and *PgXET1* amino acid sequences showed that *PgXET1* possesses a deletion of 3 amino acids (amino acids G⁹⁶, D⁹⁷ and A⁹⁸ of *Z. mays* XTH) and substitution of H with D at position 99 (Figure 3). When compared with XET sequence from *Z. mays2*, there are changes in 43 amino acids, excluding insertion and deletions (Figure 4). Deletions of amino acids (at 2 positions) and insertion of amino acids (at 3 positions) were observed. At nucleotide level there are 118 base substitutions, including 37 transitions and 81 transversions. The most frequent transversion was noticed to be G↔C occurring at 49 places. At 22 positions, substitutions of 2 to 4 bases were observed.

Recombinant expression and functional activity of XET

To verify functional activity of the cloned *PgXET1*, a recombinant *PgXET1* was produced using the *P. pastoris* heterologous expression system. The cDNA, encoding mature protein (without signal peptide), was cloned into pPICZ A expression vector in-frame with the N-terminal -factor. The protein was successfully expressed in *P. pastoris* and the purified protein was found to be about 33 kDa in molecular weight. XET activity of the purified protein was tested using fluorescence dot-spot assay by spotting *PgXET1* onto XET test paper. This resulted into the transfer of XGO-SR to the reducing end of a non-

fluorescently labeled high molecular weight xyloglucan (donor substrate) which could be observed in the form of fluorescent spot (Figure 5a). In the control reaction, fluorescently labeled non-XET acceptor substrate was used wherein no fluorescence was observed (Figure 5b). No hydrolysis activity was observed in the iodine staining assay by the recombinant *PgXET1* protein.

Phylogenetic tree

To study phylogenetic relationship between *PgXET1* and the other known XETs, a phylogenetic tree was generated using full-length protein sequences of 31 different XETs from 25 plant species (Figure 6). The evolutionary distances are presented in the units of the number of amino acid differences per site. Ambiguous positions were removed for each sequence pair and a total of 248 positions were found in the final dataset. The XET sequences were found to be clustered broadly into two major phylogenetic groups (Groups 1 and 2) with well-supported bootstrap values. The major monocot XETs were clustered into a single group. *PgXET1* was grouped (in Group 1) with XET from *Z. mays* (ZmXET2) with very strong bootstrap support (>95%). The nucleotide and amino acid sequences of the maize XET [*Z. mays2* (ADB54615)] was also aligned with those of the *PgXET1* (Figure 4). At nucleotide level, the conservation was found to be 85%. The maximum variation in the sequences was observed to be in the region

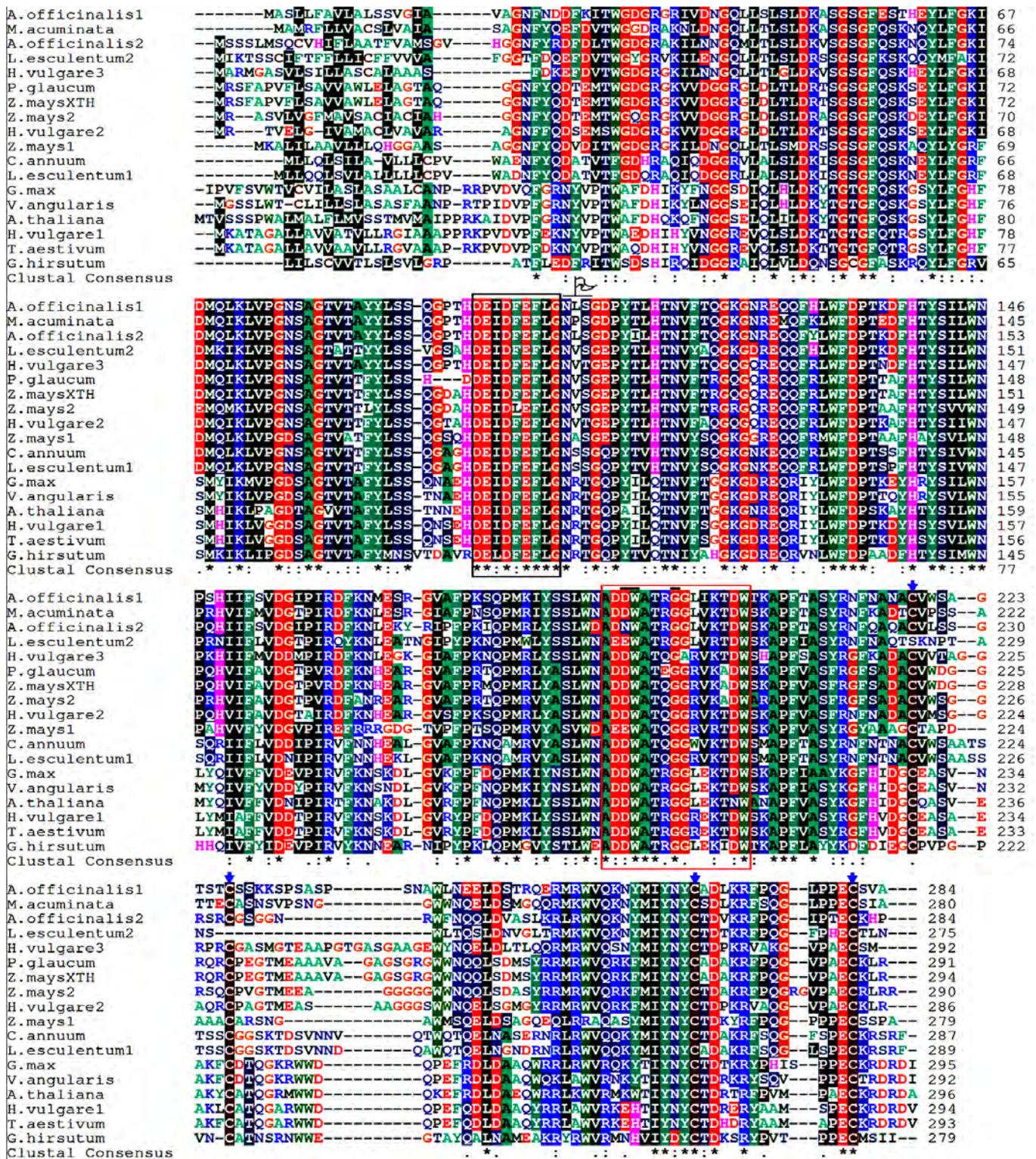


Figure 3. Alignment of deduced amino acid sequence of *PgXET1* and that of XET sequences from different plant species, including a sequence of XTH from *Z. mays* (*Z.maysXTH*, accession No. NP001151103) which showed maximum homology on BLASTn. The conserved regions from where degenerate primers were designed for cloning of the cDNA are marked by black arrows (horizontal). The conserved catalytic site is embossed in black and the immediately followed N-glycosylation site is flagged. The acceptor binding loop is embossed in red and the amino acids involved in binding of the xyloglucan substrates are marked by horizontal lines. The conserved cysteine residues (amino acids 220, 229, 274 and 288 in *P. glaucum*) are marked by vertical blue arrows.

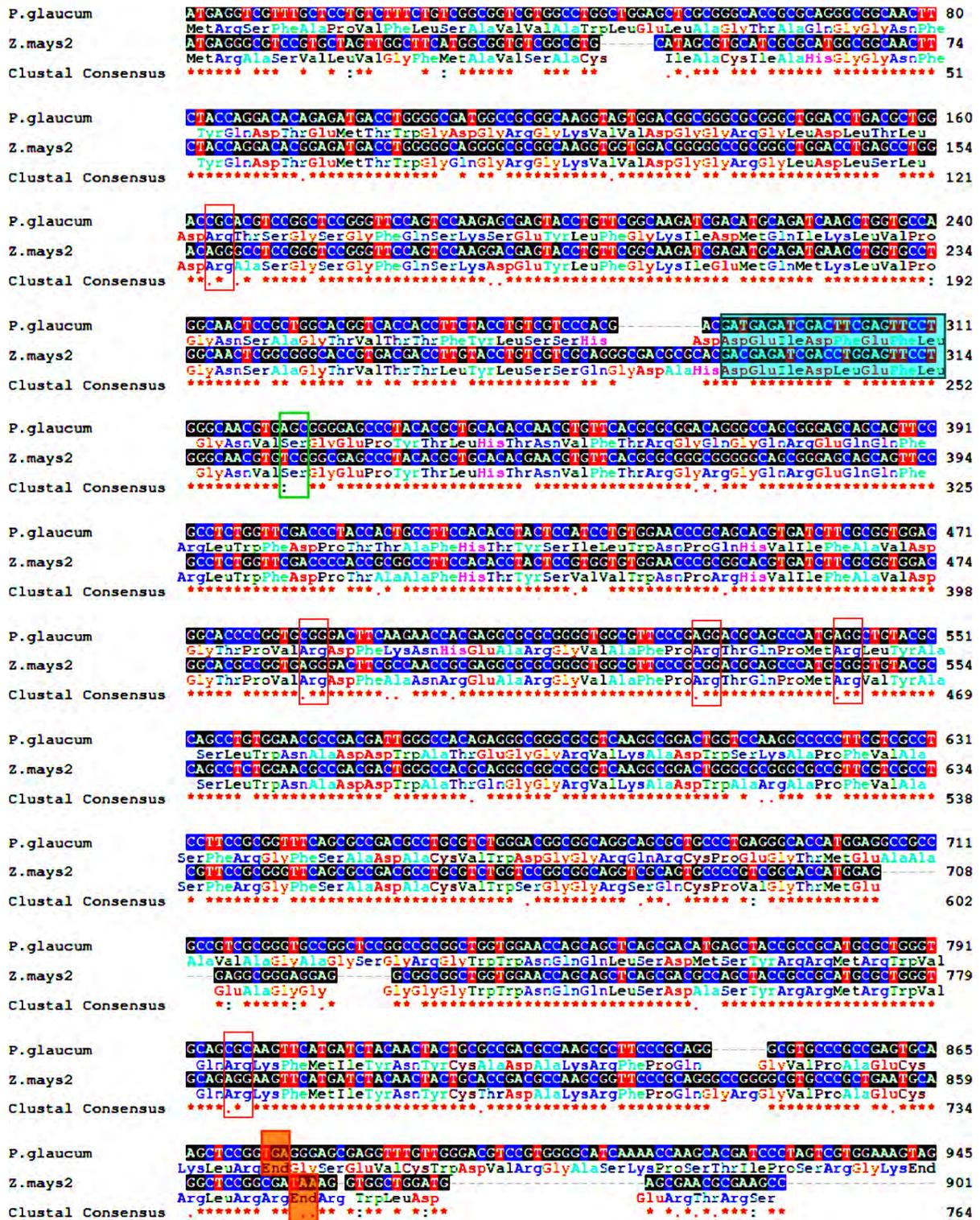


Figure 4. Alignment of XET sequences of *P. glaucum* with that of *Z. mays* (*Z.mays2*). Three letter abbreviated names of the encoded amino acids are shown in the second line. Insertions/deletions are marked by broken lines at 5 different positions (deletions at 2 places and insertion at 3 places). Five point mutations at non-wobble position of the codon which did not change amino acid (due to the degeneracy of codon) are marked with red boxes. At one position (marked with green box), even changes in all the three nucleotides of a codon, did not change the amino acid (Ser). Active site of the proteins is embossed with sky-blue; stop codons are embossed with orange.

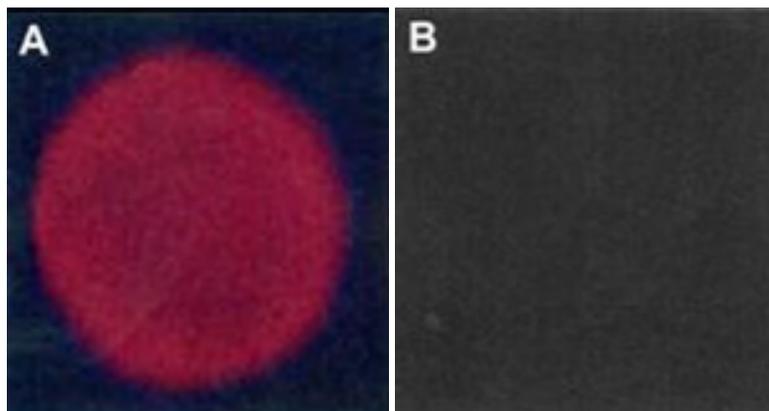


Figure 5. XET activity assay of the recombinant *PgXET1*. (A) Fluorescent spot produced by the purified recombinant *PgXET1* spotted on the xyloglucan-containing test paper, (B) no fluorescence in case of the control.

designated for signal peptide.

DISCUSSION

The role of XET in growth and development of plants is well recognized now. Dynamics of plant growth requires concomitant modifications of plant cell wall since it determines the cell shape and size. Plant canopy being responsible for trapping the solar energy, dry matter production is directly related to the amount of intercepted solar radiation by the canopy. In this study, cloning, characterization and expression analysis of the *PgXET1* was carried out to improve our understanding of XET activity in pearl millet to increase biomass production, particularly as a forage crop.

Composite cDNA, followed by the complete cDNA, from *P. glaucum* was cloned combining RT-PCR, 3'-RACE and 5'-RACE techniques. The degenerate primers designed from the conserved region of the enzyme resulted into amplification of a 297 bp fragment. Based on the sequence information, the flanking 5' and 3' sequences were successfully amplified. The composite cDNA sequence information was then used for cloning the complete cDNA (1266 bp) and confirmation of the nucleotide sequence of the *PgXET1*.

Analysis of the cloned *PgXET1* cDNA revealed that it contains ORF coding for 291 amino acids. The encoded protein has a hydrophobic amino terminus to function as a secretion signal peptide that allows localization of the protein/enzyme to the plant cell wall, its site of the enzymatic action. The detailed study of conserved amino acids and its motifs revealed that the active site (DEIDFEFLG) of the enzyme (XET) is highly conserved in *PgXET1*. This is in agreement with the earlier reports for XET from many other plants (Nishitani, 1997). A

similar active site is also found in GH16 family of glycoside hydrolases which comprises of xyloglucan endotransglucosylases, 1,3 β -glucanases and 1,3 to 1,4 glucan hydrolases-I activities in Group B and β -agarases, κ -carrageenases and 1,4 β -glucanases in Group A (Strohmeier et al., 2004). Site directed mutagenesis have earlier demonstrated that both glutamic acid residues in the conserved motif act as nucleophile and general base and are involved in the cleavage of β -glycosyl linkages (Hahn et al., 1995). Functional importance of these residues in XETs has been confirmed by different workers (Campbell and Braam, 1998; Johansson et al., 2003). The conserved sequence might be crucial for the formation and stability of enzyme itself and/or enzyme-substrate complex. The first phenylalanine residue (F) of the catalytic site has been suggested to interact with the other phenylalanine residues (Michel et al., 2001). The active site, as in most other XETs studied so far, is immediately followed by a potential site for N-linked glycosylation (N-X-T/S) (Van Sandt et al., 2006). Cysteine residues at position 220, 229, 274 and 288 are well conserved and probably involved in the formation of inter- or intra-molecular disulphide bonds to stabilize the three-dimensional structure of the C-terminal region of the enzyme. The conserved arginine, glutamic acid and aspartic acid at well defined positions are thought to play important roles in the three-dimensional structure of the active site-cleft and positioning of donor and acceptor substrates. The amino acids involved in N-glycan interactions (that is, E¹¹⁰, R¹³¹ in *PgXET1*) are generally conserved or substituted to homologous amino acids. The amino acid involved in recognition of the xyloglucan substrates (Y¹⁸³ in *PgXET1*) is conserved among the angiosperms (Figure 3). The aspartate-tryptophan (D¹⁹¹-W¹⁹²) and glycine (G¹⁹⁶) involved in binding of xyloglucan are largely conserved among the angiosperms. As

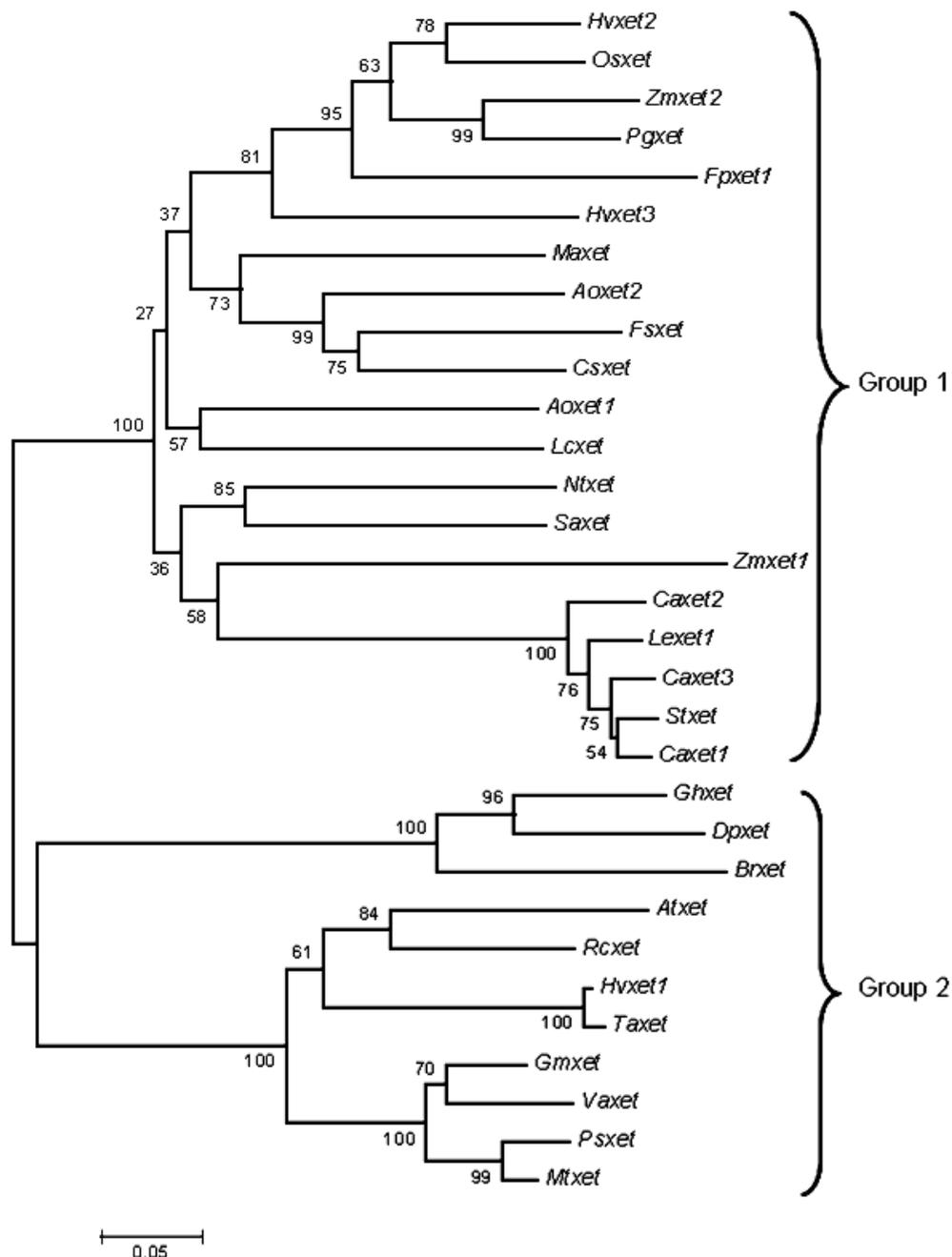


Figure 6. Phylogenetic tree of XETs from different plant species. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (5000 replicates) is shown next to the branches. Tree has been drawn to the scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. Sum of branch length = 4.08479105).

reported in the case of other XETs (Van Sandt et al., 2006), E¹¹⁰, H¹¹⁵ and F¹⁴⁰ in *PgXET1*, are important for maintaining three-dimensional structure of the active site.

Comparison of amino acid sequences of *PgXET1* with known XETs showed maximum homology with *Z. mays*

XTH (Xyloglucan endotransglucosylase/hydrolase), followed by XET from *Z. mays* (*Z. mays2*). This suggests a closer relationship between pearl millet and maize in terms of XET activities. This is again in agreement with the fact that pearl millet (*P. glaucum*) and maize (*Z.*

mays) belongs to the same sub-family Panicoideae. Minor changes in terms of deletion of 3 amino acids (amino acids 96, 97 and 98 of *Z. mays*XTH) and substitutions of Q→H and H→D at positions 95 and 99, just before the active site (Figure 3), might be responsible for the differences in the enzymatic activity of these two proteins.

Comparison of *PgXET1* amino acid sequence with that of *Z. mays2* revealed that there are changes in 43 amino acids (Figure 4). Deletions and insertion of amino acids were also observed. Most of the changes in nucleotides (96 out of 118) were due to point mutations. Almost half of the point mutations are null mutation and the remaining could cause change in 43 amino acids. To maintain the enzymatic activity of the protein, majority of the substituted amino acids are of similar physical property. Some of the changes in amino acid were due to substitution of nucleotide at either 1st or 2nd position of the codon. All the other changes in nucleotide were at the 3rd (wobble) position of the codon. Five substitutions at non-wobble position of the codon could not change amino acids because of the degeneracy of genetic code (CGU, CGA, CGC, CGG, AGA and AGG all code for Arg). Even after change of all the three nucleotides of a codon (AGC↔TCG) at one position, there was no change in amino acid (Ser), again because of the degeneracy of codons (6 degenerate codons for each of Arg, Leu and Ser). At nucleotide level, out of 118 base substitutions, 37 were transitions while 81 were found to be the transversions, owing to the more possible ways of transversions occurring at molecular level.

The recombinant *PgXET1* was expressed in yeast eukaryotic system. The signal peptide was excluded to avoid localization of the protein in the yeast cell and to harvest the protein for *in vitro* studies. As expected, the protein was found to be about 33 kDa. Fluorescence dot-spot assay showed XET activity in the purified protein due to the transfer of XGO-SR to the reducing end of a non-fluorescently labeled high molecular weight xyloglucan (donor substrate). Since iodine staining assay with the recombinant *PgXET1* did not show any hydrolysis activity, it was categorized to be an XET.

The *PgXET1* protein showed high degree of homology with ZmXET2 (accession no. ADB54615) of maize (*Z. mays*). The XET sequences were found to be clustered broadly into two major phylogenetic groups with well-supported bootstrap values. Most of the monocot XETs clustered in to a single group wherein *PgXET1* and ZmXET2 paired with very strong bootstrap support (>95%). The clustering of XETs from different plant species was found to be in agreement with their evolutionary development. Accordingly, *PgXET1* was grouped with XETs from its more closely related plants (maize, barley and rice).

The cloning of cDNA of XET, the first one from *P. glaucum* and comparative genomics by *in silico* analyses

suggest that the *PgXET1* contains high degree of conservation, particularly in the active site. The understanding of biochemical and physiological roles of XETs in plant is being enhanced by integrated approaches to gain a comprehensive knowledge of the gene networks, proteins and metabolites involved. This knowledge will lead to novel approaches for improving biomass production and drought tolerance in crop plants through genetic and metabolic engineering.

ACKNOWLEDGEMENTS

We are grateful to the Head, Division of Crop Improvement and the Director, Indian Grassland and Fodder Research Institute, Jhansi, for supports and facilities provided. S. K. acknowledges IUSSTF Research Fellowship grants from Indo-US Science and Technology Forum, New Delhi, India.

REFERENCES

- Burstin J (2000). Differential expression of two barley XET-related genes during coleoptile growth. *J. Exp. Bot.* 51: 847-852.
- Campbell P, Braam J (1998). Co- and post-translational modifications are critical for TCH4 XET activity. *Plant J.* 15: 553-561.
- Campbell P, Braam J (1999). Xyloglucan endotransglycosylases: diversity of genes, enzymes and potential wall-modifying functions. *Trends Plant Sci.* 4: 361-366.
- Carpita NC, McCann M, Griffing CR (1996). The plant extracellular matrix: news from the cell's frontier. *Plant Cell* 8: 1451-1463.
- Chou K-C, Shen HB (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* 5: e11335. doi:10.1371/journal.pone.0011335.
- Cosgrove DJ (2000). Loosening of plant cell walls by expansins. *Nature*, 407: 321-326.
- Cosgrove DJ (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6: 850-861.
- Fanutti C, Gidley MJ, Reid JSG (1993). Action of a pure xyloglucan endo-transglycosylase (formerly called xyloglucan-specific endo-1,4-β-D-glucanase) from the cotyledons of germinated nasturtium seeds. *Plant J.* 3: 691-700.
- Felsenstein J (1985). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783-791.
- Fry SC (1997). Novel 'dot-blot' assays for glycosyltransferases and glycosyl hydrolases: optimizations for xyloglucan endotransglycosylase (XET) activity. *Plant J.* 11: 1141-1150.
- Gangaiah B (2008). Agronomy of *kharif* crops. Available online at: <http://www.nsd.niscair.res.in/bitstream/123456789/540/1/Fodder> (verified: June 11, 2011).
- Hahn M, Olsen O, Politz O, Borriss R, Heinemann U (1995). Crystal structure and site-directed mutagenesis of *Bacillus macerans* endo-1,3-1,4-beta-glucanase. *J. Biol. Chem.* 270: 3080-3088.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Series* 41: 95-98.
- Harada T, Torii Y, Morita S, Onodera R, Hara Y, Yokoyama R, Nishitani K, Satoh S (2010). Cloning, characterization, and expression of xyloglucan endotransglucosylase/ hydrolase and expansin genes associated with petal growth and development during carnation flower opening. *J. Exp. Bot.* doi:10.1093/jxb/erq319.
- Johansson P, Denman S, Brumer H, Kallas AM, Henriksson H, Bergfors T, Teeri TT, Jones TA (2003). Crystallization and preliminary X-ray

- analysis of a xyloglucan endotransglycosylase from *Populus tremula* x *tremuloides*. *Acta Crystallographica*, 59: 535-537.
- McQueen-Mason S (1997). Plant cell walls and the control of growth. *Biochem. Soc. Trans.* 25: 204-214.
- Michel G, Chantalat L, Duee E, Barbeyron T, Henrissat B, Kloareg B, Dideberg O (2001). The β -carrageenase of *P. carrageenovora* features a tunnel-shaped active site: A novel insight in the evolution of Clan-B glycoside hydrolases. *Structure*, 9: 513-525.
- Nei M, Kumar S (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
- Nishitani K (1997). The role of endoxyloglucan transferase in the organization of plant cell wall. *Int. Rev. Cytol.* 173: 157-206.
- Nishitani K, Tominaga T (1992). Endo-xyloglucan transferase, a novel class of glycosyltransferase that catalyzes transfer of a segment of xyloglucan molecule to another xyloglucan molecule. *J. Biol. Chem.* 267: 21058-21064.
- Rai KN, Murty DS, Andrews DJ, Bramel-Cox PJ (1999). Genetic enhancement of pearl millet and sorghum for the semi-arid tropics of Asia Africa. *Genome*, 42: 617-628.
- Reidy B (2001). The role of cell wall-related proteins during leaf growth of *Festuca pratensis* Huds. Dissertation, Swiss Fed. Institute Technol. Zurich.
- Rose JKC, Braam J, Fry SC, Nishitani K (2002). The XTH family enzyme involved in xyloglucan endotransglucosylation and endohydrolysis: current perspectives and a new unifying nomenclature. *Plant Cell Physiol.* 43: 1421-1435.
- Saitou N, Nei M (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4: 406-425.
- Strohmeier M, Hrmova M, Fischer M, Harvey AJ, Fincher GB, Pleiss J (2004). Molecular modeling of family GH16 glycoside hydrolases: potential roles for xyloglucan transglucosylases/ hydrolases in cell wall modification in the poaceae. *Protein Sci.* 13: 3200-3213.
- Suda CNK, Buckeridge MS, Giorgini JF (2003). Cell wall hydrolases in the seeds of *Euphorbia heterophylla* L. during germination and early seedling development. *Braz. J. Plant Physiol.* 15: 135-143.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* (In Press).
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.
- Van Sandt VST, Guisez Y, Verbelen JP, Vissenberg K (2006). Analysis of a xyloglucan endotransglycosylase/hydrolase (XTH) from the lycopodiophyte *Selaginella kraussiana* suggests that XTH sequence characteristics and function are highly conserved during the evolution of vascular plants. *J. Exp. Bot.* 57: 2909-2922.
- Wu Y, Cosgrove DJ (2000). Adaptation of roots to low water potentials by changes in cell wall extensibility and cell wall proteins. *J. Exp. Bot.* 51: 1543-1553.
- Wu Y, Sharp RE, Durachko DM, Cosgrove DJ (1996). Growth maintenance of the maize primary root at low water potentials involves increases in cell wall extensibility, expansin activity and wall susceptibility to expansins. *Plant Physiol.* 111: 765-772.
- Wu Y, Thorne ET, Sharp RE, Cosgrove DJ (2001). Modification of expansin transcript levels in the maize primary root at low water potentials. *Plant Physiol.* 126: 1471-1479.
- Xu W, Campbell P, Vargheese AV, Braam J (1996). The Arabidopsis XET-related gene family: Hormonal and environmental regulation of expression. *Plant J.* 9: 879-889.
- Yokoyama R, Nishitani K (2001). A comprehensive expression analysis of all members of a gene family encoding cell-wall enzymes allowed us to predict cis-regulatory regions involved in cell-wall construction in specific organs of Arabidopsis. *Plant Cell Physiol.* 42: 1025-1033.