American Journal of Bioinformatics

Original Research Paper

Bioinformatics Analysis of Xyloglucan Endotransglycosylase/Hydrolase (XTH) Gene from Developing Xylem of a Tropical Timber Tree *Neolamarckia Cadamba*

[1]Tiong Shing Yiing, [1]Ho Wei Seng, [2]Pang Shek Ling and [1]Ismail Jusoh

[1]*Forest Genomics and Informatics Laboratory (fGiL), Department of Molecular Biology,Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak*
[2]*Applied Forest Science and Industry Development (AFSID), Sarawak Forestry Corporation, 93250 Kuching, Sarawak*

Article history
Received 2014-07-31
Revised 2014-10-27
Accepted 2014-10-31

Corresponding Author:
Ho Wei Seng,
Forest Genomics and Informatics Laboratory (fGiL), Department of Molecular Biology, Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak
Email: wsho@frst.unimas.my

Abstract: This study reported the isolation and *in silico* characterization of a full-length Xyloglucan endotransglycosylase/Hydrolase (XTH) cDNA from *Neolamarckia cadamba*, an important tropical light hardwood plantation tree species. XTH is considered as a key agent to regulate cell wall expansion and is believed to be responsible for the incorporation of newly synthesised xyloglucan into the wall matrix. The full-length of *NcXTH* was firstly predicted using the XTH singletons from the NcdbEST through contig mapping approach. Further validation and confirmation were conducted by amplifying the full-length XTH cDNA using RT-PCR approach. Two full-length XTH cDNAs, namely *NcXTH1* (JX134619) and *NcXTH2* (JX134620) were discovered and the nucleotide sequences were 893 and 1,024 bp in length, respectively. The open reading frames for *NcXTH1* and *NcXTH2* were 858 and 915 bp, respectively. Results predicted that *NcXTH1* and *NcXTH2* proteins carry out XET activity but they are from different XTH family members. This full-length *NcXTH* cDNA can serve as good candidate genes in association genetics study which leads to Gene-Assisted Selection (GAS) in the *N. cadamba* tree breeding programme.

Keywords: *Neolamarckia Cadamba*, Xyloglucan Endotransglycosylase/Hydrolase (XTH), Wood Formation, Contig Mapping, Protein Structure Prediction, Phylogenetics

Introduction

Wood formation or also known as xylogenesis is an ordered and complex developmental process in plants. It involves cell division, cell expansion, secondary wall deposition, lignification and programmed cell death. Most of the enzymes that involved in cell wall biopolymers synthesis are under the Carbohydrate Active enzymes (CAZymes) family. These include Glycosyltransferases (GTs), Glycoside Hydrolases (GHs), Polysaccharide Lyases (PLs) and various Carbohydrate Esterases (CEs) (Geisler-Lee *et al*., 2006). Xyloglucan endotransglycosylase/Hydrolase (XTH) is one of the glycosyltransferase members. XTH is considered as a key agent to regulate cell wall expansion and is believed to be responsible for the incorporation of newly synthesised xyloglucan into the wall matrix. In order for cell wall to expend, cross-linked or connections of microfibrils need to be broken. XTH is able to cut a xyloglucan chain and rejoin the reducing end to another xyloglucan molecule (xyloglucan endotransglycosylase, XET action) or to water molecule (xyloglucan endotranshydrolase, XEH action) (Fry *et al*., 1992; Darley *et al*., 2001; Rose *et al*., 2002).

A few mechanisms have been proposed in XET action. XTHs may join the newly synthesised xyloglucans into a larger xyloglucan polymer chains before integrated with existing xyloglucans in the cell wall layer (Campbell and Braam, 1999a). Transient polysaccharide-enzyme complex is probably formed between XTH and xyloglucan as the intermediate before being transferred and joined to another xyloglucan (Sulová *et al*., 1998). XTHs may also be involved in shifting the size of xyloglucans by XET action that allows cellulose microfibrils to move apart and/or past one another driven by tugor pressure (Fry *et al*., 1992; Talbott and Pickard, 1994). When there is no new xyloglucan released and supplied to the wall, XET activity could lead to rearrangement of existing xyloglucans or degradation of deposited xyloglucans (Campbell and Braam, 1999a). Studies also suggested that XET activities may play an important role in fruit development. Decrease in *XTH* genes expression during fruit ripening suggest that XET action might contributes to fruit softening (Miedes and Lorences, 2009).

XTHs are family genes. *XTH* is reported as one of the key gene families in GH16 that are involved in cell walls modification (Ye *et al*., 2012). Expression studies of CAZymes in poplar shows that XTHs are one of the most highly expressed cell wall enzyme (Geisler-Lee *et al*., 2006). The evolution of *XTH* family gene through gene duplication and divergence has brought to the multi-members being reported: 33 *XTH* genes in *Arabidopsis* (Yokoyama and Nishitani, 2001); 41 in *populus* (Geisler-Lee *et al*., 2006); 29 in rice (Yokoyama *et al*., 2004; Yokoyama and Nishitani, 2004). Researchers have divided these large family members into a few subfamilies according to the gene structures and expression: Three (Rose *et al*., 2002) or four (Saladié *et al*., 2006) subfamilies in *Arabidopsis*; two main groups in rice (Yokoyama *et al*., 2004); three (Geisler-Lee *et al*., 2006) or four (Ye *et al*., 2012) groups in poplar; three major clusters in tomato and kiwi (Atkinson *et al*., 2009).

XTH proteins are predicted to have several structural features in common: A hydrophobic amino terminus which probably functions as a signal peptide to direct the protein to the cell wall; a highly conserved DEIDFEFLG domain that acts as the catalytic site for both transferase and hydrolase activities; an *N*-linked glycosylation consensus site which its function for XET activity still remain unclear; and pairs (four) of Cysteine (C) residues in the carboxyl-terminal region that might form disulphide bridges (Okazawa *et al*., 1993; Campbell and Braam, 1999b). Although many *XTH* genes and proteins were discovered, their detail functions in vascular tissues and in the formation of secondary walls are less well understood.

To date, there are considerable amounts of full-length XTH cDNA being published in NCBI but no such information available for *Neolamarckia cadamba* trees. *N. cadamba* or locally known as kelampayan belongs to the family of Rubiaceae. It has been selected as one of the fast growing plantation species for planted forest development in Sarawak (Tchin *et al*., 2012; Lai *et al*., 2013; Tiong *et al*., 2014a; 2014b; Ho *et al*., 2014). The state government of Sarawak has introduced the Forest (Planted Forest) Rules (1997) to encourage the development of commercial planted forests and has set a target of 1.0 million hectares for forest plantations to be established by 2020. It is estimated that 42 million of high quality seedlings are required for the annual planting programme. *N. cadamba* is a large, deciduous and fast growing tree that gives early economic returns within 8-10 years. Under normal conditions, it attains a height of 17 m and diameter of 25 cm at breast height (dbh) within 9 years. It is a lightweight hardwood with a density of 290-560 kg/m^3 at 15% moisture content (Joker, 2000). It is one of the best sources of raw material for the plywood industry, besides pulp and paper production. *N. cadamba* can also be used as a shade tree for dipterocarp line planting, whilst its leaves and bark have medical applications. The dried bark can be used to relieve fever and as a tonic, whereas a leaf extract can serve as a mouth wash (WAC, 2004). *N. cadamba* also has high potential to be utilized as one of the renewable resource of raw materials for bioenergy production such as cellulosic biofuels in the near future.

Hence, the objectives of this study were: (i) To obtain the full-length *XTH* cDNA sequences through contig mapping approach by using *XTH* singletons from the Kelampayan tree transcriptome database (NcdbEST) and (ii) to *in silico* characterize the *XTH* genes from *N. cadamba*. The full-length XTH cDNA discovered can serve as good candidate gene for association genetics study in *N. cadamba* to detect the potential genetic variants underlying the common and complex adaptive traits.

Materials and Methods

Hypothetical Full-Length XTH cDNAs Assembly from EST Singletons

Singletons of *XTH* gene were selected from the Kelampayan tree transcriptome database (NcdbEST) (Ho *et al*., 2014). The *XTH* singletons were blast again NCBI database to search for sequence homology and binding position on the respective gene. Subsequently, the singletons were grouped according to the alignment score and position on gene. Singletons which have overlapping fragment were then identified and jointed together to form the full-length sequences via contig mapping approach. Two hypothetical cDNAs (*XTH1* and *XTH2*) were used to design primer pairs for full-length *XTH* cDNAs amplification by using Primer Premier 5 software (PREMIER Biosoft International, USA). The oligonucleotide primers used for *XTH1* were NcXTH1-F (5'-ACAATGGCTTCTCATTTGAACT-3') and NcXTH1-R (5'-TTTGGCTCCTCTCAGATCG-3') and *XTH2* were NcXTH2-F (5'-CTTCTGATTCATCAATGGCTTC-3') and XTH2-R (5'-CATAGAGTTCATGTCCAGTGCA-3').

RNA Isolation and RT-PCR Amplification

Total RNA was isolated from the developing xylem tissues of *N. cadamba* using RNeasy® Midi Kit (QIAGEN GmbH, Germany) with modification. Total RNA was then reverse transcript into cDNA by using Ready-To-Go You-Prime First-Strand Beads (GE Healthcare, UK). RT-PCR amplification was carried out in a total reaction volume of 25 µL containing 1 x Advantage 2 PCR buffer (Clontech, USA), 1.5 mM $MgCl_2$, 0.2 mM of dNTPs, 10 pmol of primer pair, 1 x Advantage 2 Polymerase Mix (Clontech, USA) and 1.0 µL of cDNA. The thermal cycling profile was programmed at 94°C for 2 min as the initial denaturation step, followed by 35 cycles of 30 sec denaturation step at 94°C, 45 sec at 57°C for annealing and 1 min extension at 72°C. The full-length PCR amplicons were purified from agarose gel by using QIAquick® Gel Extraction Kit (QIAGEN, Germany). Purified PCR product was ligated into pGEM®-T Easy Vector System (Promega, USA) and transformed into competent cells, *Escherichia coli* JM 109. The recombinant plasmids were isolated and purified using Wizard® *Plus* SV Minipreps DNA Purification System (Promega, USA) according to the manufacture's protocol. After verification, the purified plasmids were sent for sequencing in both forward and reverse direction. The sequencing reactions were performed by using ABI Prism™ Bigdye ™ terminator cycle sequencing Ready reaction kit V.3.1 (Applied Biosystems, USA) and analysed on a ABI 3730 XL capillary DNA sequence (Applied Biosystems, USA).

Sequence Analysis of Full-Length XTH Genes

The full-length *NcXTH1* and *NcXTH2* cDNAs that had been sequenced were manually edited using Chromas Lite version 2.01 programme to remove the vector sequence. Sequence homology search for *NcXTH1* and *NcXTH2* were performed against GenBank non-redundant nucleotide sequence using the NCBI Basic Local Alignment Search Tool (BLAST) server. Expert Protion Analysis System (ExPASy) translate tool (Gasteiger *et al*., 2003) was use to change the nucleotide sequence into amino acid sequence and to find the open reading frame (ORF). Amino acid sequences of *NcXTH1* and *NcXTH2* were used to predict motifs through multiple alignments with protein sequence of *XTH* genes from other species using EBI-ClustalW2 multiple alignment tool and CLC Main Workbench version 5.0 software (CLC bio, Denmark). Signal peptide was predicted using SignalP 4.0 Server (Petersen *et al*., 2011). The transmembrane helices in protein were predicted using TMHMM Server version 2.0 (Krogh *et al*., 2001).

Three-Dimensional (3D) Protein Structure Prediction

Secondary or 3D structure of NcXTH1 and NcXTH2 were predicted using intensive mode of Phyre2 server by homology modeling approach (Kelley and Sternberg, 2009). Ligand binding site prediction server (3DLigandSite) (Wass *et al*., 2010) was used to predict the sugar binding sites and active sites of NcXTH1 and NcXTH2 in secondary structure. Predicted 3D structures of NcXTH1 and NcXTH2 were viewed using Jmol, an open-source Java viewer for chemical structures in 3D with features for chemicals, crystals, materials and biomolecules (Jmol, 2012). Vector Alignment Search Tool (VAST) provided by NCBI was used to identify similar 3D structures in the Molecular Modeling Database (MMDM) and compared with predicted XTH protein structures. Sequence similarity (% Id) for the parts of the protein that have been superimposed and SCORE (the VAST structure-similarity score) that reflects the quality of superimposed elements were recorded (Panchenko and Madej, 2004).

Phylogenetic Analysis of Full-Length XTH Genes

The full-length *XTH* genes were obtained from NCBI nucleotide database and the GenBank accession numbers were recorded. All of the *A. thaliana XTH* genes were obtained from the *Arabidopsis* Information Resource (TAIR) database (http://arabidopsis.org) using the gene models accession. Protein sequence of selected genes was aligned using EBI-ClustalW2 multiple alignment tool and Neighbour-Joining (NJ) tree was generated using MEGA version 5 software (Tamura *et al*., 2011).

Results

Hypothetical Full-Length XTH cDNAs Assembly from Singletons

A total of eight *XTH* singletons were identified from the NcdbEST. The *XTH* singletons (ranging from 300 bp to 762 bp) showed high similarity (scored up to 83% identity) when BLAST with *XTH* genes of other species, such as *A. hemsleyana* (EU494954), *A. deliciosa* (EU494953), *P. tremula* x *P. tremuloides* (EF151160) and *V. angularis* (EF599289) (Table 1). From the BLAST results, singletons Ncdx016F05, Ncdx053B07, Ncdx104G02 and Ncdx106G11 shared the maximum identity with *XTH* gene (EU494954) from a kiwi species (*A. hemsleyana*). Singleton Ncdx106G11 was longer than Ncdx016F05 and the alignment of these two sequences scored at 100 and therefore three singletons, Ncdx106G11, Ncdx053B07 and Ncdx104G02, were selected for contig mapping to produce a hypothetical XTH1 cDNA sequence. Another hypothetical full-length *XTH2* cDNA was contig mapped from another two singletons, Ncdx099D12 and Ncdx044A08.

NcXTH1 and NcXTH2 cDNA Sequences Analysis

The assembled hypothetical full-length *XTH* cDNA sequences were used to design two full-length primer pairs for *XTH* genes amplification in *N. cadamba*. The amplified full-length *XTH* cDNAs from two full-length primer pairs were named as *NcXTH1* and *NcXTH2*, respectively. The sequencing result of *NcXTH1* cDNA was aligned with its assembled

DOI: 10.3844/ajbsp.2014.1.16

hypothetical full-length cDNA sequence (*XTH1*) as shown in Fig. 1. *NcXTH2* cDNA also showed high similarity with its hypothetical full-length cDNA sequence (*XTH2*) in the alignment as shown in Fig. 2. BLAST analysis of *NcXTH1* and *NcXTH2* cDNAs against the NCBI database is shown in Table 2.

NcXTH1 cDNA that contained 858 bp of the coding sequence (cds) had translated 855 bp of its open reading frame into 285 amino acids. For *NcXTH2* cDNA, a 915 bp of cds containing 912 bp ORF was translated into 304 amino acids. Amino acid sequences of NcXTH1 and NcXTH2 were aligned as shown in Fig. 3. The alignment score of 45.0 suggested that NcXTH1 and NcXTH2 might be two different *XTH* members. However, both *XTH* genes were predicted to be involved in similar biochemical functions. Figure 4 shows their shared common conserved features for XTH proteins when aligned with a few closely related *XTH* genes as listed in Table 2. The diagrammatic representation of NcXTH1 and NcXTH2 proteins are shown in Fig. 5 and 6, respectively.

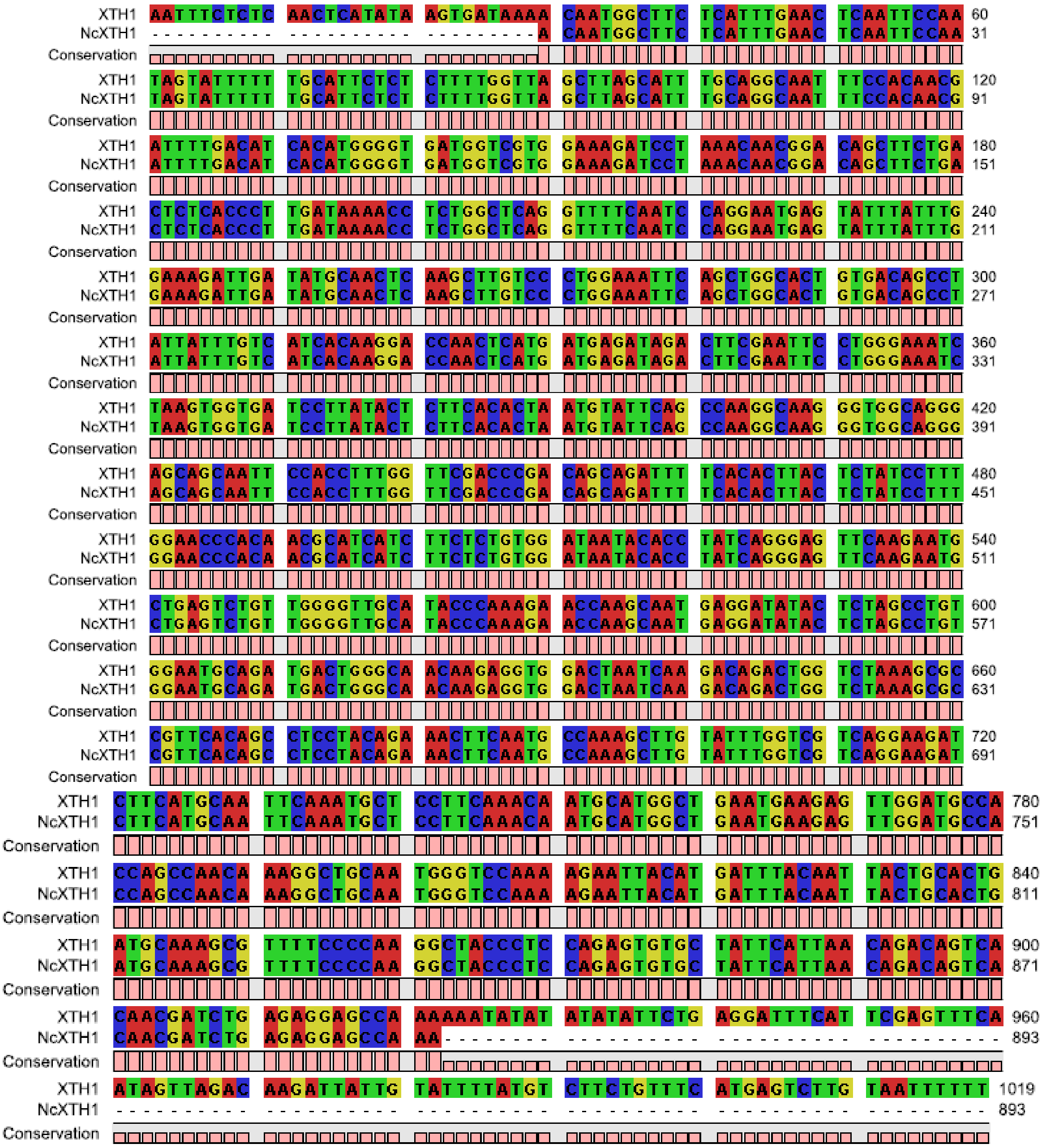


Fig. 1. Sequence alignment of full-length NcXTH1 cDNA with XTH1 hypothetical full-length cDNA

DOI: 10.3844/ajbsp.2014.1.16

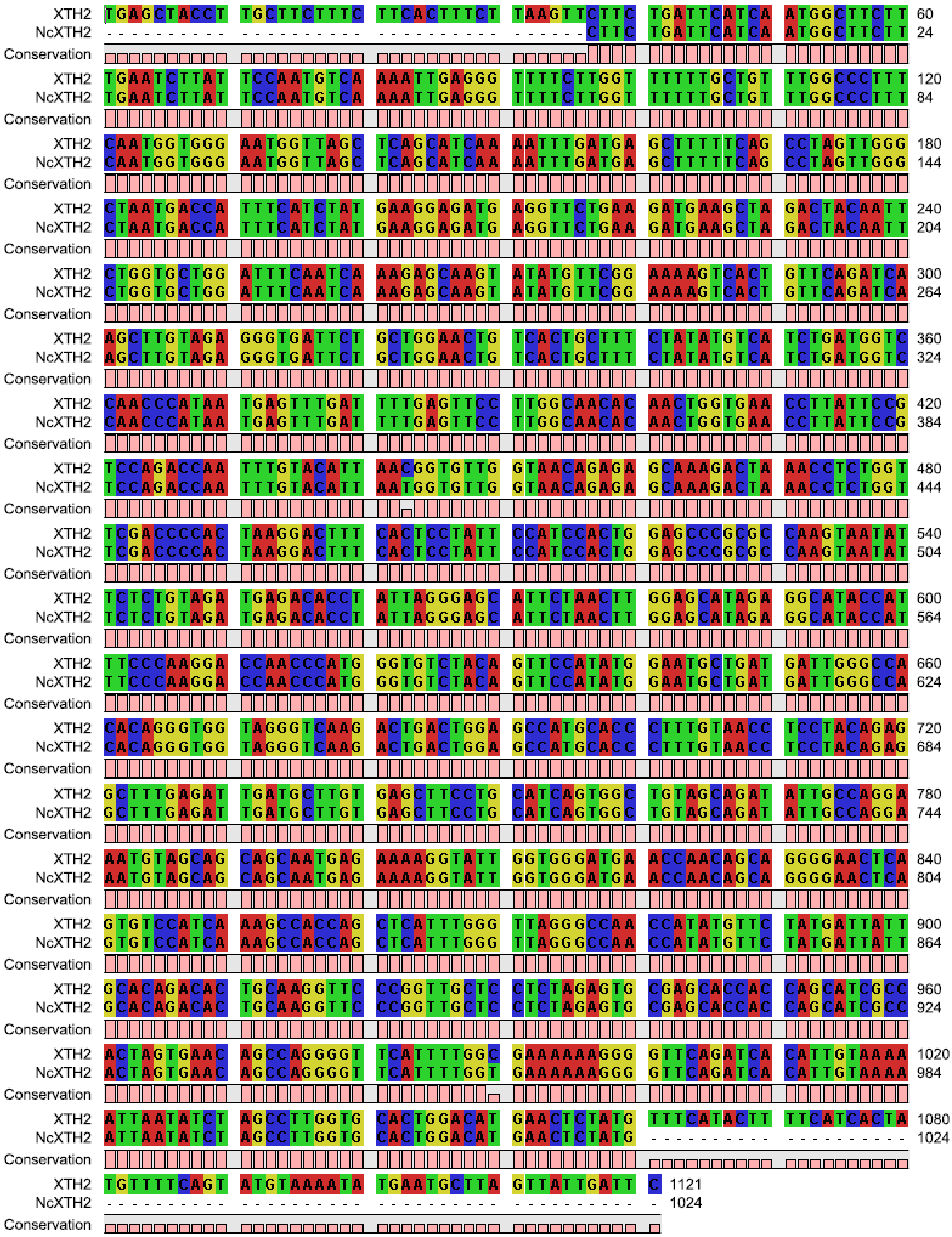


Fig. 2. Sequence alignment of full-length NcXTH2 cDNA with XTH2 hypothetical full-length cDNA. The alignment region was almost identical with only two differing nucleotides

DOI: 10.3844/ajbsp.2014.1.16

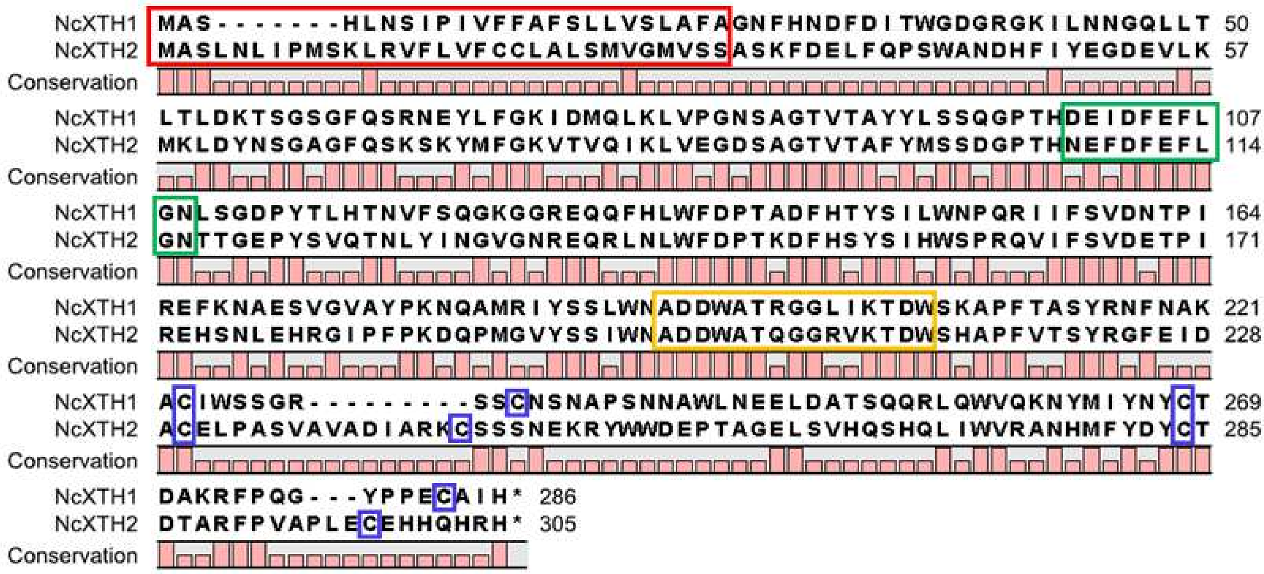
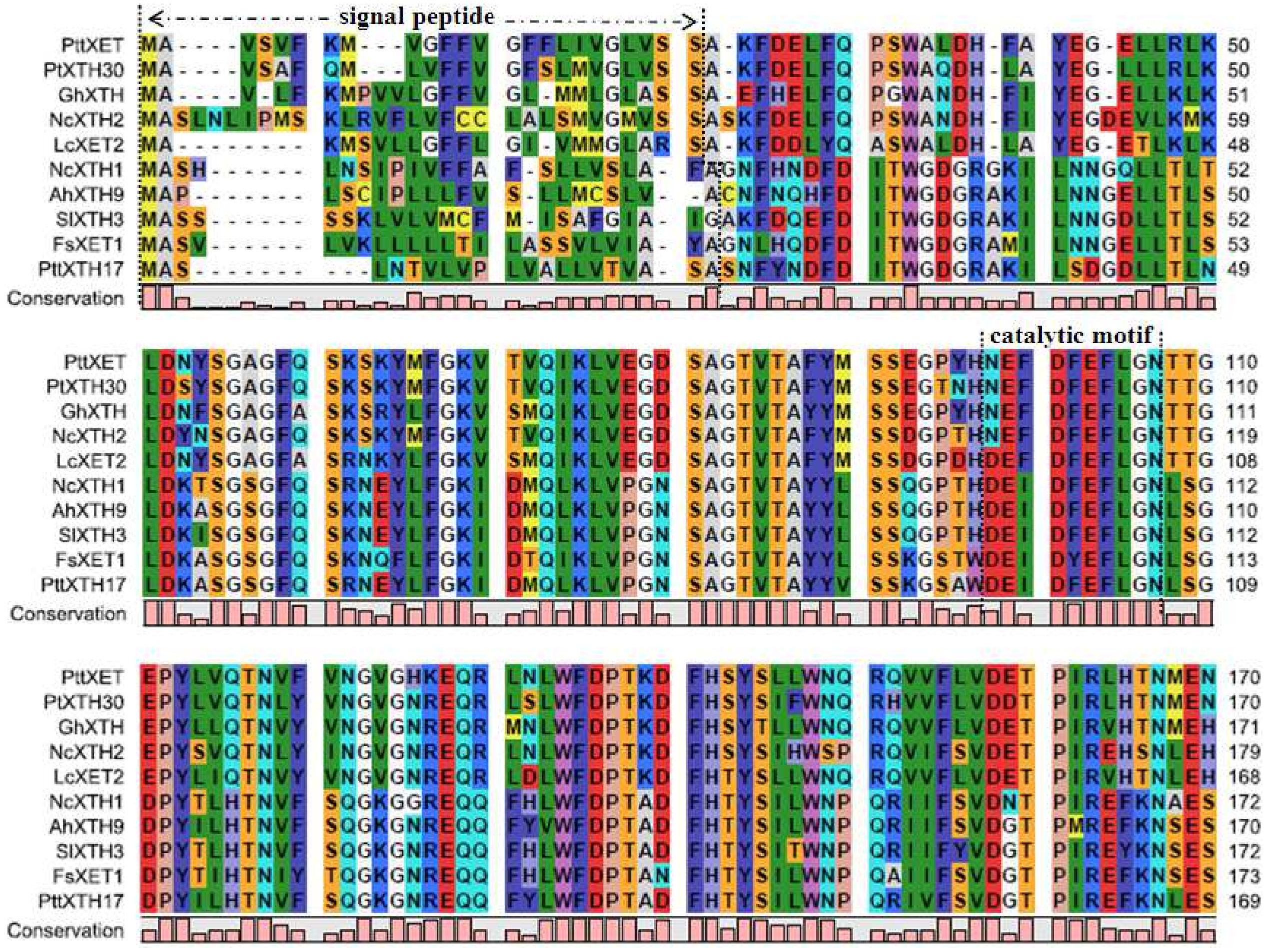


Fig. 3. NcXTH1 and NcXTH2 amino acid sequences alignment. Signal peptide (red box), catalytic domain (green box) and N-glycosylation site (orange box) and cysteines (blue boxes) were predicted in both NcXTH1 and NcXTH2 protein sequences



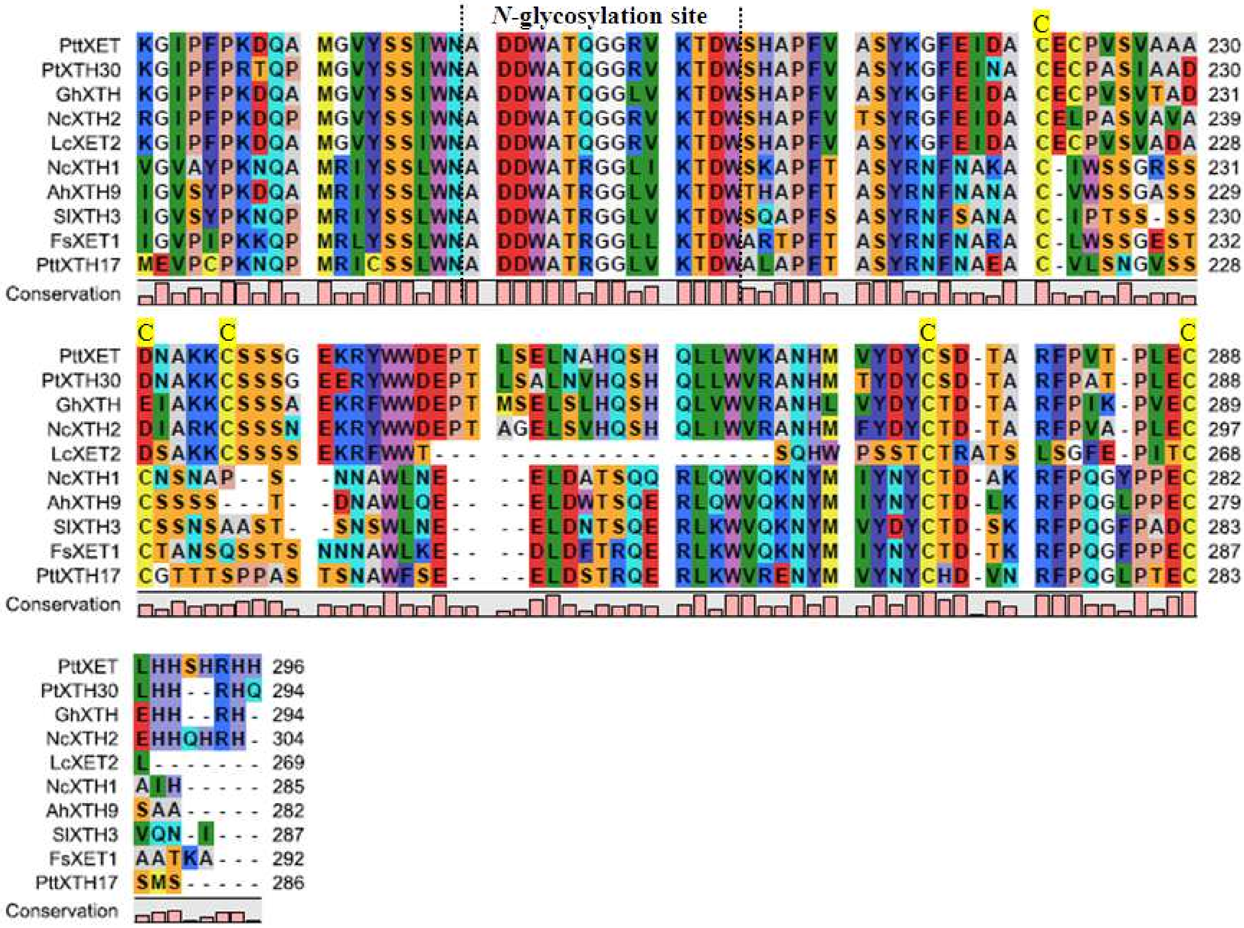


Fig. 4. Amino acid sequence alignment of NcXTH1 and NcXTH2 with XTH protein sequences of *A. hemsleyana* (AhXTH9; EU494954), *F. sylvatica* (FsXET1; AJ130885), *G. hirsutum* (GhXTH; JN968478), *L. chinensis* (LcXET2; DQ995514), *P. tremula* (PtXTH30; EF194057), *P. tremula* x *P. tremuloides* (PttXET; EF151160 and PttXTH17; EF194056) and *S. lycopercium* (SlXTH3; AY497476). Conserved regions and catalytic domain are shown as indicated on top of each alignment line

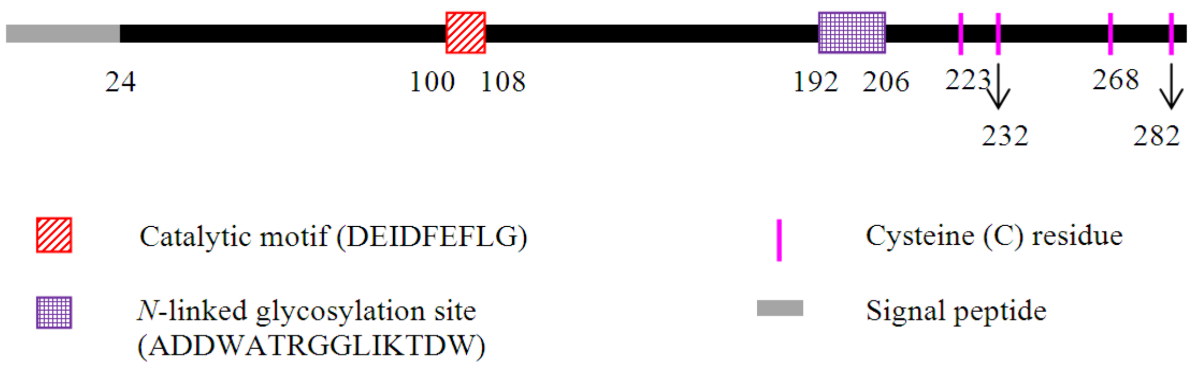


Fig. 5. A diagrammatic representation of NcXTH1 protein (285 amino acids) with the common structural features of XTH proteins shown by the coloured boxes and lines and its respective location in the amino acid sequence shown by the numbers

DOI: 10.3844/ajbsp.2014.1.16

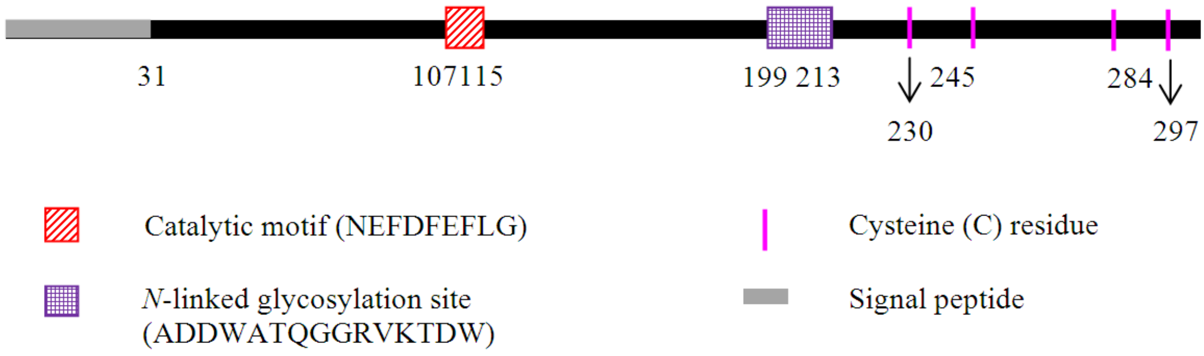


Fig. 6. A diagrammatic representation of NcXTH2 protein (304 amino acids) with the common structural features of XTH proteins shown by the coloured boxes and lines and its respective location in the amino acid sequence shown by the numbers

Table 1. BLAST results of *XTH* singletons that showed the highest similarity with other species *XTH* genes

		BLAST results		
Singleton	Length (bp)	Identity (%)	E value	Accession no.
Ncdx016F05	506	79	$3e^{-88}$	EU494954
Ncdx053B07	681	82	$2e^{-162}$	EU494954
Ncdx104G02	557	82	$1e^{-106}$	EU494954
Ncdx076F06	300	83	$5e^{-25}$	EU494953
Ncdx044A08	417	77	$4e^{-41}$	EF151160
Ncdx099D12	762	79	$3e^{-154}$	EF151160
Ncdx106G11	679	79	$6e^{-130}$	EU494954
Ncdx033C10	630	80	$4e^{-75}$	EF599289

Table 2. BLAST sequence homology analysis of full-length *NcXTH1* and *NcXTH2* cDNAs against NCBI GenBank database nucleotide collection (nr/nt)

Organism	Accession number	Coverage (%)	Identity (%)	E value
NcXTH1				
Actinidia hemsleyana	EU494954	86	80	0
Fagus sylvatica	AJ130885	90	77	$4e^{-177}$
Lycopersicum esculentum/ Solanum lycopercium	AY497476	85	78	$7e^{-175}$
Populus tremula x *Populus tremuloides*	EF194056	86	76	$4e^{-158}$
Asparagus officinalis	AF223420	87	74	$1e^{-146}$
Rosa hybrid	AB428380	87	75	$4e^{-146}$
Malus domestica	EU494967	83	75	$2e^{-145}$
Populus tremula x *Populus tremuloides*	EF194054	82	76	$2e^{-144}$
NcXTH2				
Populus tremula x *Populus tremuloides*	EF151160	82	79	0
Gossypium hirsutum	JN968478	80	78	0
Litchi chinensis	DQ995514	85	76	$7e^{-177}$
Populus tremula	EF194057	83	94	$8e^{-176}$
Malus domestica	EU494964	83	76	$3e^{-168}$
Actinidia eriantha	EU494948	86	74	$5e^{-153}$
Rosa hybrid	AB428381	81	75	$2e^{-152}$
Actinidia deliciosa	EU494949	80	75	$1e^{-147}$

DOI: 10.3844/ajbsp.2014.1.16

3-D structure Prediction of NcXTH1 and NcXTH2

Full-length protein sequences of NcXTH1 and NcXTH2 with 285 and 304 amino acids, respectively, were used to predict the three-Dimensional (3D) secondary protein structures using intensive mode of Phyre2 server (Kelley and Sternberg, 2009). Figure 7 shows the final models generated with front view (a and c) and side view (b and d) for NcXTH1 and NcXTH2 protein, respectively. The confidence level, which is the probability of the query sequence and the template sequence are homologous, were high for both NcXTH1 and NcXTH2 secondary structures (more than 90% confidence level). The concave face of both NcXTH1 and NcXTH2 was predicted using 3DLigandSite server (Wass *et al*., 2010) as shown in Fig. 8a and brespectively.

NcXTH1 and NcXTH2 Phylogenetic Analysis

A total of 79 *XTH* genes from various species were obtained from NCBI database and renamed according to the systematic nomenclature of *XTH* genes (Rose *et al*., 2002). The designated abbreviations were: *Actinidia chinensis* (*Ac*); *Actinidia deliciosa* (*Ad*); *Actinidia eriantha* (*Ae*); *Actinidia hemsleyana* (*Ah*); *Asparagus officinalis* (*Ao*); *Arabidopsis thaliana* (*At*); *Fagus sylvatica* (*Fs*); *Gossypium hirsutum* (*Gh*); *Malus domestica* (*Md*); *Litchi chinensis* (*Lc*); *Lycopersicon esculentum* (*Le*); *Populus tremula* (*Pt*); *Populus tremula* x *Populus tremuloides* (*Ptt*); *Ricinis communis* (*Rc*); *Rosa hybrid* (*Rh*); *Shorea parvifolia* (*Sp*). Amino acid sequences of all 79 *XTH* genes were aligned with both NcXTH1 and NcXTH2 amino acid sequences to generate a neighbour-joining (NJ) phylogenetic tree as shown in Fig. 9.

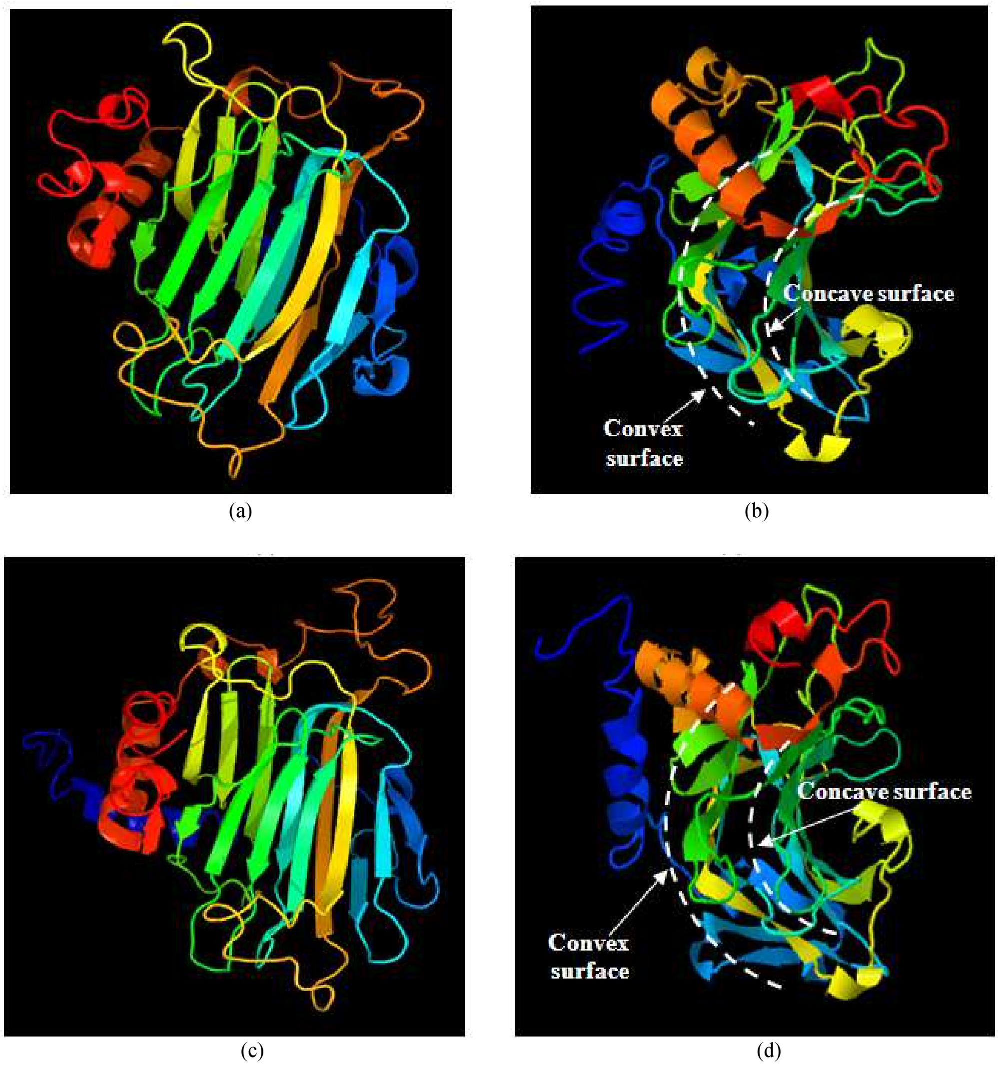
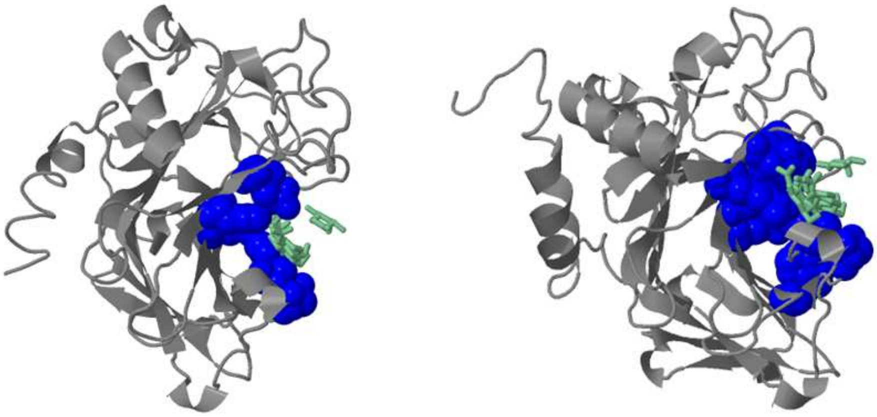


Fig. 7. NcXTH1 predicted 3D protein structure with the front view (a and c) and the side view (b and d) of NcXTH1 (a-b) and NcXTH2 (c-d). N- to C-terminal were coloured from blue to red



(a) (b)

Fig. 8. Side view of NcXTH1 (a) and NcXTH2 (b) secondary structure showing the substrate binding site at concave face indicated by the blue particles. Light green wires refer to the heterogens-non-standard residues, prosthetic groups, inhibitors, solvents (except water) or ions that are unknown to the Worldwide Protein Data Bank (wwPDB)

Discussion

Eight *XTH* singletons were selected from the NcdbEST. This database was generated through high-throughput 5'-EST sequencing of cDNA clones derived from developing xylem tissues of *N. cadamba*. It consists of a total of 10,368 EST with 6,622 showed high quality EST sequences (Ho *et al*., 2014). These partial cDNA sequences were blasted, aligned, joined through contig-mapping approach to produce a longer hypothetical *XTH* sequences. A full-length *XTH* sequence named as *XTH1* was hypothetically mapped with the length of 1,019 bp and 855 bp Open Reading Frame (ORF) from three singletons Ncdx106G11, Ncdx053B07 and Ncdx104G02. *XTH1* cDNA scored the best hit with kiwi (*A. hemsleyana*) *XTH9* complete cds, as expected, which covered 75% of the sequences with 80% sequence similarity and E-value of 0. *XTH* genes of other species that also showed good coverage (>70%) and identity (>75%) with *XTH2* included European beech (*F. sylvatica*), *Populus*, *Asparagus*, apple (*M. domestica*) and rose (*Rosa hybrid*).

Another hypothetical full-length *XTH* cDNA named as *XTH2* was also mapped from two singletons Ncdx099D12 and Ncdx044A08. The consensus sequence *XTH2* was 1,121 bp long with 922 bp ORF. The sequence homology analysis of *XTH2* showed the best hit with hybrid *Populus* species (EF151160) with the sequence similarity up to 78% and covered 75% coverage of the sequence. Complete coding sequence (cds) of *P. tremula XTH-30* (EF194057) showed the highest sequence similarity (94%) with 76% coverage. *XTH* genes of lychee (*L. chinensis*), apple (*M. domestica*), tomato (*L. esculentum*), kiwi (*A. eriantha* and *A. deliciosa*) and rose (*Rosa hybrid*) also showed high nucleotide sequence coverage (>70%) and identity (>75%) with *XTH2*.

The two hypothetical *XTH* cDNAs were then used to design primer pairs for full-length *XTH* cDNAs amplification. The alignment result showed that the *NcXTH1* was 100% identical to *XTH1* with no nucleotide polymorphism or gap in the alignment (Fig. 1). This further confirmed that the targeted *XTH* gene had been successfully cloned and sequenced. The total length of *NcXTH1* cDNA sequence (GenBank accession number: JX134619) was 893 bp. A start codon was found at nucleotides 4-6 and a stop codon was found at nucleotides 859-861. *NcXTH1* contains 858 bp of coding sequence (inclusive stop codon) with 3 bp of incomplete 5'-Untranslated Region (UTR) and 32 bp of incomplete 3'-UTR. *NcXTH2* cDNA (GenBank accession number: JX134620) with a total length of 1,024 bp also showed high similarity with its hypothetical full-length cDNA sequence *XTH2*. This result confirmed that the designed primer pair had amplified the targeted gene and the correct insert was cloned. A start codon (ATG) and a stop codon (TAG) were found at nucleotides 15-17 and 927-929, respectively (Fig. 2). *NcXTH2* had a longer coding region (915 bp) compared to *NcXTH1* (858 bp) with the incomplete 14 bp 5'-UTR and 95 bp 3'-UTR. BLAST analysis of *NcXTH1* and *NcXTH2* cDNAs against the NCBI database supported that these two genes as *XTH* family members with the identity values of up to 80 and 94%, respectively with poplar or aspen (*Populus*), kiwi (*A. hemsleyana*), tomato (*L. esculentum*), apple (*M. domestica*), rose (*Rosa hybrid*) and some other species (Table 2). *NcXTH1* showed the highest sequence similarity (80%) with kiwi fruit (*A. hemsleyana*), which covered 86% of its full-length *XTH* mRNA. On the other hand, *NcXTH2* showed a very high sequence similarity (94%) with *P. tremula*, which covered 83% of the nucleotide sequence of XTH30 protein.

DOI: 10.3844/ajbsp.2014.1.16

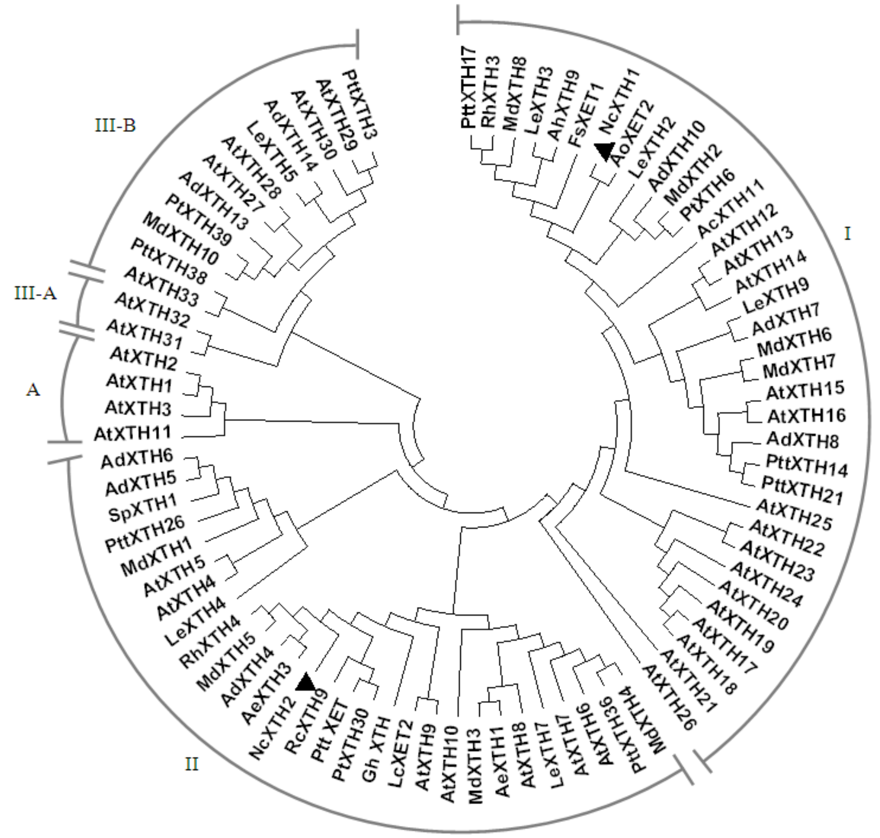


Fig. 9. Neighbour-joining phylogenetic tree generated from the amino acid sequences alignment of NcXTH1 and NcXTH2 with 79 XTH amino acid sequences of other species using the ClustalW and MEGA5 software; GenBank accession number of XTHs used are: NcXTH1 (JX134619); NcXTH2 (JX134620); AcXTH11 (EU494956); AdXTH4 (EU494949); AdXTH5 (EU494950); AdXTH6 (EU494951); AdXTH7 (EU494952); AdXTH8 (EU494953); AdXTH10 (EU494955); AdXTH13 (EU494958); AdXTH14 (EU494959); AeXTH1 (EU494946); AeXTH3 (EU494948); AhXTH9 (EU494954); AoXET2 (AF223420); AtXTH1 (AT4G13080); AtXTH2 (AT4G13090); AtXTH3 (AT3G2505); AtXTH4 (AT2G0685); AtXTH5 (AT5G13870); AtXTH6 (AT5G65730); AtXTH7 (AT4G37800); AtXTH8 (AT1G11545); AtXTH9 (AT4G03210); AtXTH10 (AT2G14620); AtXTH11 (AT3G48580); AtXTH12 (AT5G57530); AtXTH13 (AT5G57540); AtXTH14 (AT4G25820); AtXTH15 (AT4G14130); AtXTH16 (AT3G23730); AtXTH17 (AT1G65310); AtXTH18 (AT4G30280); AtXTH19 (AT4G30290); AtXTH20 (AT5G48070); AtXTH21 (AT2G18800); AtXTH22 (AT5G57560); AtXTH23 (AT4G25810); AtXTH24 (AT4G30270); AtXTH25 (AT5G57550); AtXTH26 (AT4G28850); AtXTH27 (AT2G01850); AtXTH28 (AT1G14720); AtXTH29 (AT4G18990); AtXTH30 (AT1G32170); AtXTH31 (AT3G44990); AtXTH32 (AT2G36870); AtXTH33 (AT1G10550); FsXET1 (AJ130885); GhXTH (JN968478); LcXET2 (DQ995514); LeXTH2 (AF176776); LeXTH3 (AY497476); LeXTH4 (AF186777); LeXTH5 (AY497475); LeXTH7 (AY497478); LeXTH9 (AY497479); MdXTH1 (EU494960); MdXTH2 (EU494961); MdXTH3 (EU494962); MdXTH4 (EU494963); MdXTH5 (EU494964); MdXTH6 (EU494965); MdXTH7 (EU494966); MdXTH8 (EU494967); MdXTH10 (EU494969); SpXTH1 (GQ338421); PtXTH6 (EF194049); PtXTH30 (EF194057); PtXTH39 (EF194055); PttXTH (EF151160); PttXTH3 (EF194053); PttXTH14 (EF194054); PttXTH17 (EF194056); PttXTH21 (EF194058); PttXTH26 (EF194046); PttXTH36 (EF194050); PttXTH38 (EF194047); RcXTH9 (XM_002522655); RhXTH3 (AB428380); RhXTH4 (AB428381). I, II, III-A and III-B indicate four main XTH subfamilies and A is the predicted ancestral group. NcXTH1 and NcXTH2 were clustered in Group II and Group I, respectively

A predicted signal peptide was found in both NcXTH1 and NcXTH2 at the first 24 and first 31 amino acids, respectively in the translated protein sequences (Fig. 3). The outputs of SignalP 4.0 server also predicted that both NcXTH1 and NcXTH2 are secretory proteins with the D-mean scored at a high value (0.904 and 0.825, respectively). This is supported by a few studies that showed the release of XTH proteins by secretory vesicles into the cell wall layers to carry out its function (Driouich *et al*., 1993; White *et al*., 1993; Zhong and Ye, 2009). The signal peptide amino acid sequence of NcXTH1 and NcXTH2 showed a greater sequence variation compared to the mature protein (without signal peptide) region. Analysis of *Arabidopsis* XETs also showed that the putative signal peptide regions have greater amino acid sequence differences (Xu *et al.*, 1996). Although the amino termini differ significantly, all of them are hydrophobic and contain the functional signal peptide (Campbell and Braam, 1999b) which has been shown to involve in direct secretion of XTH into the apoplast (Yokoyama and Nishitani, 2001).

The conserved domain which acts as the catalytic site for all XTHs was found in both NcXTH1 and NcXTH2. NcXTH1 possesses the major catalytic motif sequence DEIDFEFLG whereas NcXTH2 has a slightly different catalytic sequence at the first and third amino acids as underlined (NEFDFEFLG). The minor amino acid difference in this catalytic domain does not affect its function because the active site (ExDxE, where x can be any amino acid) is always conserved. Campbell and Braam (1999b) had compared a few *Arabidopsis* XTH catalytic domains and found out that the third residue, Isoleucine (I) may also be replaced by another hydrophobic residues, either Leucine (L) or Valine (V) and the first phenylalanine (F) (fifth residue) may be substituted by I. These changes are predicted to have no effect on the cleavage of β-1,4-glycosyl linkages because the apolar and uncharged nature of the residues are still maintained. However, the change of the first glutamate residue (abbreviated as Glu or E) which acts as the active site has shown to inactivate the protein (Campbell and Braam, 1998).

A putative *N*-linked glycosylation site was also found in both NcXTH1 and NcXTH2 at nucleotides 192-206 and nucleotides 199-213, respectively, with three amino acid differences as underlined (ADDWATR/QGGL/R I/VKTDW). This potential site is probably recognized by the plant cell glycosylation machinery (Campbell and Braam, 1999b). Although the importance of the *N*-glycosylation site still remains unclear, it was shown to have significant influence on XET activity (Campbell and Braam, 1999a). The removal of N-linked glycosylation has eliminated 98% of the XET activities (Campbell and Braam (1998). A total of four highly conserved cysteine (C) residues were found in the carboxyl termini of NcXTH1 and NcXTH2 amino acid sequences. Each pair of cysteine residue has the potential to form a disulphide bond, either inter-or intra-molecularly (Campbell and Braam, 1999b). Disulphide bond formation and reshuffling between cysteine residues has an important role in co- and post-translational protein modification, which contributes to the protein folding pattern and its stability (Huppa and Ploegh, 1998). Campbell and Braam (1998) found that Trans-Cinnamate 4-Hydroxylase (TC4H) gene encodes an XET activity, where the reduction of disulphide bond (s) on this gene caused significant decreases in XET activity. Therefore, this bonding is believed to be important for full XET activity and essential for the stability of the most active conformation of the enzyme.

At the 100% confidence level, excluding the signal peptide region, both NcXTH1 and NcXTH2 secondary structures demonstrated the highest scores in sequence similarity with percentage identity (% Id) of 43.3 and 40.1%, respectively and in secondary structure similarity (SCORE: 29.5 and 28.9, respectively) with two crystal structures: *P. tremula* x *P. tremuloides* XET16A (PttXET16A) [Protein Data Bank Identity (PDB ID): 1UMZ] and Glycoside Hydrolase family 16 (GH16) endoxyloglucanase TmNXG1 (PDB ID: 2VH9). This means that the 3D protein model of NcXTH1 and NcXTH2 were reliable. Secondary structure of the signal peptide protein at 5'-end of both NcXTH1 and NcXTH2 were modelled by *ab initio* approach with a much lower confidence level (Xu *et al.*, 1996) since this region has very low sequence similarity compared to those in the database. The overall structures of NcXTH1 and NcXTH2 were similar to other enzymes in GH16, which they share the β-jelly-roll fold with two β-sheets aligned in a sandwich like manner (Strohmeier *et al.*, 2004). The first 3D structure solved for family GH16 was from a *Bacillus* 1,3-1,4-β-glucanase which was determined by x-ray crystallography (Keitel *et al.*, 1993). They found that the two antiparallel β-sandwich consisted of one concave and one convex surface. These structures were also found in both NcXTH1 and NcXTH2 as shown in Fig. 7b and d, respectively. Similar structures were also observed in the first eukaryotic XTH crystal structure of *Populus* hybrid PttXET16A (Johansson *et al.*, 2004).

An α-helix and a short β-sheet found at the C-terminal in NcXTH1 and NcXTH2 (coloured in red as shown in Fig. 7) were predicted to be responsible for XET activity. This notable structural feature was also observed in PttXET16A, but not in other family GH16 enzymes having the C terminus located after the final β-strand of the β-sheet (Johansson *et al.*, 2004). The importance of the C-terminal α-helix in NcXTH1 and NcXTH2 also can be revealed by its high structure confidence level on the predicted α-helix structure. Another α-helix located at N-terminus (blue coloured) was observed in both NcXTH1 and NcXTH2 with the α-helix of NcXTH2 being longer than the one in NcXTH1. However, the role or functions of this structural feature at the N-terminus has yet to be discovered and was not observed in PttXET16A. Therefore, the structural confidence of N-terminus α-helix was very low.

The concave face of both NcXTH1 and NcXTH2 contains sugar binding sites as predicted using 3DLigandSite server (Wass *et al.*, 2010) as shown in Fig. 8a and b, respectively. The center β-sheet of the concave face has been suggested to be the location where catalytic or active side chains were offset from (Ståhlberg *et al.*, 1996; Johansson *et al.*, 2004) and most of the family GH16 members share a general active site motif of ExDxE (Michel *et al*., 2001). The active sites of NcXTH1 and NcXTH2 on β-sheet were also supported by the clustering of XTH proteins into subgroup 2 of family GH16 in Clan-B. Michel *et al*. (2001) suggested that Clan-B glycoside hydrolases fall into two subgroups, with most of the GH members having catalytic machinery held by an ancestral β-bulge and were grouped into subgroup 1. Another group (subgroup 2) with the active sites held by a regular β-strand, consists of only XTHs and 1,3-1,4-glucanases (lichenases) and were suggested to be evolved from family GH16.

All of the selected 81 *XTH* genes, inclusive *NcXTH1* and *NcXTH2* (Fig. 9) were classified into main subfamily I, II and III as discussed by a few research papers (Geisler-Lee *et al*., 2006). However, some reported that the divergence between group I and group II was not apparent and therefore classified these two groups into a big subfamily I/II (Yokoyama *et al*., 2004; Baumann *et al*., 2007; Michailidis *et al*., 2009). XTHs in subfamily I and II were reported to be the most highly expressed cell wall-related CAZymes involved in the XET activity (Geisler-Lee *et al*., 2006; Ye *et al*., 2012). In this study, groups I and II were assigned separately as discussed by Nishikubo *et al*. (2011; Ye *et al*., 2012) using *XTH* genes of a timber species, *Populus*. Subfamily III was also further clustered into group III-A and III-B, according to their activity (Geisler-Lee *et al*., 2006; Baumann *et al*., 2007; Ye *et al*., 2012).

Recently, an "ancestral group" (indicated as group A in Fig. 9) was introduced as these sequences are closest to a bacterial β-1,3-1,4-glucanase (Baumann *et al*., 2007; Michailidis *et al*., 2009; Ye *et al*., 2012). Sequence analysis showed that these sequences (AtXTH1, AtXTH2, AtXTH3 and AtXTH11) clade between group I/II and group III as the intermediate, possibly the ancestral (Baumann *et al*., 2007). In the putative ancestral group, researchers suggested that group I was most likely to occur before the rise of group II and group III-B subsequently, in separate events (Eklöf and Brumer, 2010).

NcXTH1 and other 34 XTH members were grouped under subfamily II, which comprise the most members among three of the subfamilies. An update by Eklöf and Brumer (2010) showed that most of the dicotyledons (inclusive *P. trichocarpa* and *A. thaliana*) and monocotyledons have the highest number of *XTH* genes in Group II. Ye *et al*. (2012) suggested that the abundance of Group II XTH family members might be due to tandem duplication as its major mechanism. In this study, three sister locus pairs of *Populus* (PtXTH17 and PtXTH18; PtXTH12 and PtXTH42; PtXTH24 and PtXTH10) from subfamily II were identified in the predicted chromosomal distribution diagram. NcXTH1 and other XTH proteins in group II were predicted to carry out XET activity as various species under this group were proven to be involved in XET processes (Eklöf and Brumer, 2010), including AtXTH14, AtXTH21, AtXTH24, AtXTH26 and MdXTH2.

NcXTH2 was found to be grouped under subfamily I with 27 other selected *XTH* genes. In contrast with group II, group I XTH members seem to be dominated by genome wide and segmental duplications (Ye *et al*., 2012). The relative abundance of group I *XTH* genes found in bryophyte (*Physcomitrella patens)* and *locophyte (Selaginella moellendorffii*) genomics suggested that group I is likely to be the original *XTH* gene product subfamily (Eklöf and Brumer, 2010). The study of *Populus PttXET16A* gene, a member of subfamily I (not shown in this study), was one of the most abundant XTH isoforms in *Populus* that demonstrated high XET activity expression in xylem and phloem fibers during secondary cell wall formation (Bourquin *et al*., 2002). Heterologous expression of *XTH* genes in group I of various species (including kiwi *AdXTH5*) also have been shown to exhibit XET activity exclusively (Eklöf and Brumer, 2010). Therefore, NcXTH1 was predicted to carry out endotransglycosylase activity rather than hydrolysis.

Subfamily III consists of 14 selected *XTH* genes from various species. Historical group III (Campbell and Braam, 1999b) can be further divided into two clades, group III-A and III-B, according to its role and activities. This was supported by the by sequence analysis, structural differences and catalytic measurements done recently by a group of researchers to find the evidence for XET and XEH activities (Baumann *et al*., 2007). They showed that only AtXTH31 and AtXTH32 (Group III-A) are predicted to have xyloglucanase (enzymatic hydrolysis) activity according to their three-dimensional structure analysis on the extension of loop 2 when compared to endo-xyloglucanase Tm-NXG1 from nasturtium (*Tropaeolum majus*). Truncation of this loop statistically decreases or diminishes hydrolytic activity of Tm-NXG1. Therefore, the variation in length of this loop 2 was predicted as the determinant of XET or Hydrolytic (XEH) activity. In contrast, none of the enzymes from group III-B, to their knowledge, possesses this hydrolytic activity. AtXTH27 (EXGT-A3) from *Arabidopsis* (Campbell and Braam, 1999b), LeXTH5 (SlXTH5) from tomato (Saladié *et al*., 2006) in Group III-B also have been shown to be predominantly or exclusively involved in XET activity. Discovery of XEH activity-related genes and knowledge about XEH activities are still very limited due to the XET activity-related genes are more abundance. Hence, most of the studies carried out on XTH proteins were related to XET activities.

XTHs that carry out XET activity will catalyze the endolytic cleavage of a cross-linking xyloglucan polymer to allow cell expansion and then transfer the newly generated end to another xyloglucan polymer to restore the primary cell wall structure (Smith and Fry, 1991; Campbell and Braam, 1999a; Thompson and Fry, 2001). A recent study suggested that this protein may also be involved in reinforcing the connection between primary and secondary cell wall layers during the early phase of secondary cell wall deposition (Bourquin *et al*., 2002). Therefore, NcXTH1 and NcXTH2 were found and believed to be abundant in developing xylem tissues of *N. cadamba*.

Conclusion

Two full-length *XTH* cDNAs, *NcXTH1* (JX134619) and *NcXTH2* (JX134620), were successfully isolated and characterized from *N. cadamba*. The conserved genetic structural features were identifed in both *XTH* genes and hence, confirmed the enzymatic role of *NcXTH1* and *NcXTH2*. Further phylogenetic analysis also predicted that these two genes are abundantly distributed and involved in XET activity, especially in the secondary cell walls. The identified *XTH* genes in this study will provide a useful resource for identifying molecular mechanisms controlling wood formation in future and will also be candidates for association genetic studies aiming at the production of high value forests (Thumma *et al*. 2005; Ho *et al*., 2011; Tchin *et al*., 2011; 2012; Tiong *et al*., 2014b; Tan *et al*., 2014). Furthermore, the detailed understanding on the regulation of XTH gene could provide a greater impact on the design of future genetic improvement strategies in the production of better quality wood that is typically present in the secondary walls of xylem in *N. cadamba*.

Acknowledgement

The researchers would like to thank all the lab assistants and foresters involved in this research project for their field assistance in sample collection. This study is part of the joint Industry-University Partnership Programme, a research programme funded by the Sarawak Forestry Corporation (SFC), Sarawak Timber Association (STA) and Universiti Malaysia Sarawak under grant no. 02(DPI09)832/2012(1), RACE/a(2)/884/2012(02) and GL(F07)/06/2013/STA-UNIMAS(06).

Author's Contributions

All authors equally contributed in this work.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Atkinson, R.G., S.L. Johnson, Y.K. Yauk, N.N. Sharma and R. Schröder, 2009. Analysis of Xyloglucan endotransglycosylase/Hydrolase (XTH) gene families in kiwifruit and apple. Postharvest Biol. Technol., 51: 149-157. DOI: 10.1016/j.postharvbio.2008.06.014

Baumann, M.J., J.M. Eklöf, G. Michel, A.S. Kallas and T.T. Teeri *et al*., 2007. Structural evidence for the evolution of xyloglucanase activity from xyloglucan endotransglycosylases: Biological implications for cell wall metabolism. Plant Cell, 19: 1947-1963. DOI: 10.1105/tpc.107.051391

Bourquin, V., N. Nishikubo, H. Abe, H. Brurner and S. Denman *et al*., 2002. Xyloglucan endotransglycosylases have a function during the formation of secondary cell walls of vascular tissues. Plant Cell, 14: 3073-3088. DOI: 10.1105/tpc.007773

Campbell, P. and J. Braam, 1998. Co- and/or post-translational modifications are critical for TCH4 XET activity. Plant J., 15: 553-561. DOI: 10.1046/j.1365-313X.1998.00239.x

Campbell, P. and J. Braam, 1999a. *In vitro* activities for four xyloglucan endotransglycosylase from *Arabidopsis*. Plant J., 18: 371-382. DOI: 10.1046/j.1365-313X.1999.00459.x

Campbell, P. and J. Braam, 1999b. Xyloglucan endotransglycosylases: Diversity of genes, enzymes and potential wall-modifying functions. Trends Plants Sci., 4: 361-366. DOI: 10.1016/S1360-1385(99)01468-5

Darley, C.P., A.M. Forrester and S.J. McQueen-Mason, 2001. The molecular basis of plant cell wall extension. Plant Molecular Biol., 47: 179-195. DOI: 10.1023/A:1010687600670

Driouich, A., G.F. Zhang and L.A. Staehelin, 1993. Effect of Brefeldin a on the structure of the golgi apparatus and on the synthesis and secretion of proteins and polysaccharides in sycamore maple (*Acer pseudoplatanus*) suspension-cultured cells. Plant Physiol., 101: 1363-1373. DOI: 10.1104/pp.101.4.1363

Eklöf, J. and H. Brumer, 2010. The XTH gene family: An update on enzyme structure, function and phylogeny in xyloglucan remodelling. Plant Physiol., 153: 456-466. DOI: 10.1104/pp.110

Fry, S.C., R.C. Smith, K.F. Renwick, D.J. Martin and S.K. Hodge *et al*., 1992. Xyloglucan endotransglycosylase, a new wall-lossening enzyme activity from plants. Biochem. J., 282: 821-828.

Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi and R.D. Appel *et al*., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res., 31: 3784-3788. DOI: 10.1093/nar/gkg563

Geisler-Lee, J., M. Geisler, P.M. Coutinho, B. Segerman and N. Nishikubo et al., 2006. Poplar carbohydrate-active enzymes: Gene identification and expression analyses. Plant Physiol., 140: 946-962. DOI: 10.1104/pp.105.072652

Ho, W.S., S.L. Pang and A. Julaihi, 2014. Identification and analysis of expressed sequence tags present in xylem tissues of kelampayan (*Neolamarckia cadamba* (Roxb.) Bosser). Physiol. Molecular Biol. Plants, 20: 393-397. DOI: 10.1007/s12298-014-0230-x

Ho, W.S., S.L. Pang, P. Lau and J. Ismail, 2011. Sequence variation in the cellulose synthase (*SpCesA1*) gene from *Shorea parvifolia* ssp. *parvifolia* mother trees. J. Tropical Agric. Sci., 34: 323-329.

Huppa, J.B. and H.L. Ploegh, 1998. The eS-sence of -SH in the ER. Cell, 92: 145-148. DOI: 10.1016/S0092-8674(00)80907-1

Jmol, 2012. Jmol: An open-source Java viewer for chemical structures in 3D. Jmol.

Johansson, P., H. Brumer III, M.J. Baumann, A.M. Kallas and H. Henriksson *et al*., 2004. Crystal structures of a poplar xyloglucan endotransglycosylase reveal details of transglycosylation acceptor binding. Plant Cell, 16: 874-886. DOI: 10.1105/tpc.020065

Joker, D., 2000. Seed Leaflet: *Neolamarckia cadamba* (Roxb.) Danida Forest Seed Centre, Bosser. Denmark.

Keitel, T., O. Simon, R. Borriss and U. Heinemann, 1993. Molecular and active-site structure of a *Bacillus* 1,3-1,.4-beta-glucanase. Proc. Nat Acad. Sci., 90: 5287-5291.

Kelley, L.A. and M.J.E. Sternberg, 2009. Protein structure prediction on the web: A case study using the Phyre server. Nature Protocols, 4: 363-371. DOI: 10.1038/nprot.2009.2

Krogh, A., B. Larsson, G. von Heijne and E.L.L. Sonnhammer, 2001. Predicting transmembrane protein topology with a hidden markov model: Application to complete genomes. J. Molecular Biol., 305: 567-580. DOI: 10.1006/jmbi.2000.4315

Lai, P.S., W.S. Ho and S.L. Pang, 2013. Development, characterization and cross-species transferability of Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) markers derived from kelampayan tree transcriptome. Biotechnology, 12: 225-235. DOI: 10.3923/biotech.2013.225.235

Michailidis, G., A. Argiriou, N. Darzentas and A. Tsaftaris, 2009. Analysis of Xyloglycan endotransglycosylase/Hydrolase (XTH) genes from allotetraploid (*Gossypium hirsutum*) cotton and its diploid progenitors expressed during fiber elongation. J. Plant Physiol., 166: 403-416. DOI: 10.1016/j.jplph.2008.06.013

Michel, G., L. Chantalat, E. Duee, T. Barbeyron and B. Henrissat *et al*., 2001. The kappa-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. DOI: 10.1016/S0969-2126(01)00612-8

Miedes, E. and E.P. Lorences, 2009. Xyloglucan endotransglycosylase/hydrolases (XTHs) during tomato fruit growth and ripening. J. Plant Physiol., 166: 489-498. DOI:10.1016/j.jplph.2008.07.003

Nishikubo, N., J. Takahashi, A.A. Roos, M. Derba-Maceluch and K. Piens *et al*., 2011. Xyloglucan endotransglycosylase-mediated xyloglucan rearrangements in developing wood of hybrid aspen. Plant Physiol., 155: 399-413.

Okazawa, K., Y. Sato, T. Nakagawa, K. Asada and L. Kato *et al*., 1993. Molecular cloning and cDNA sequencing of endoxyloglucan transferase, a novel class of glycosyltransferase that mediates molecular grafting between matrix polysaccharides in plant cell walls. J. Biol. Chem., 268: 25364-25368.

Panchenko, A.R. and T. Madej, 2004. Analysis of protein homology by assessing the (dis)similarity in protein loop regions. Proteins, 57: 539-547. DOI: 10.1002/prot.20237

Petersen, T.N., S. Brunak, G. von Heijne and H. Nielsen, 2011. SignalP 4.0: Discriminating signal peptides from transmembrane regions. Nature Meth., 8: 785-786. DOI: 10.1038/nmeth.1701

Rose, J.K.C., J. Braam, S.C. Fry and K. Nishitani, 2002. The XTH family of enzymes involved in xyloglucan endotransglycosylation and endohydrolysis: Current perspectives and a new unifying nomenclature. Plant Cell Physiol., 43: 1421-1435. DOI: 10.1093/pcp/pcf171

Saladié, M., J.K.C. Rose, D.J. Cosgrove and C. Catalá, 2006. Characterization of a new Xyloglucan endotransglycosylase/Hydrolase (XTH) from ripening tomato fruit and implication for the diverse modes of enzymic action. Plant J., 47: 282-295. DOI: 10.1111/j.1365-313X.2006.02784.x

Smith, R.C. and S.C. Fry, 1991. Endotransglycosylation of xyloglucans in plants cell suspension culture. Biochem. J., 279: 529-535.

Ståhlberg, J., C. Divne, A. Koivula, K. Piens and M. Claeyssens *et al*., 1996. Activity studies and crystal structures of catalytically deficient mutants of cellobiohydrolase I from *Trichoderma reesei*. J. Molecular Biol., 264: 337-349. DOI: 10.1006/jmbi.1996.0644

Strohmeier, M., M. Hrmova, M. Fischer, A.J. Harvey and G.B. Fincher *et al*., 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. DOI: 10.1110/ps.04828404

Sulová, Z., M. Takácová, N.M. Steele, S.C. Fry and V. Farkas, 1998. Xyloglucan endotransglycosylase: Evidence for the existence of a relatively stable glycosyl-enzyme intermediate. Biochem. J., 330: 1475-1480.

Talbott, L.D. and B.G. Pickard, 1994. Differential changes in size distribution of xyloglucan in the cell walls of gravitropically responding *Pisum sativum* epicotyls. Plant Physiol., 106: 755-61. DOI: 10.1104/pp.106.2.755

Tamura, K., D. Peterson, N. Peterson, G. Stecher and M. Nei *et al*., 2011. MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol. Biol. Evol., 28: 2731-9. DOI: 10.1093/molbev/msr121

Tan, C.J., W.S. Ho and S.L. Pang, 2014. Resequencing and nucleotide variation of Sucrose Synthase (*SuSy*) gene in a tropical timber tree *Neolamarckia macrophylla*. Int. J. Scientific Technol. Res., 3: 135-140.

Tchin, B.L., W.S. Ho, S.L. Pang and J. Ismail, 2011. Gene-associated Single Nucleotide Polymorphism (SNP) in Cinnamate 4-Hydroxylase (*C4H*) and Cinnamy Alcohol Dehydrogenase (*CAD*) genes from *Acacia mangium* superbulk trees. Biotechnology, 10: 303-315. DOI: 10.3923/biotech.2011.303.315

Tchin, B.L., W.S. Ho, S.L. Pang and J. Ismail, 2012. Association genetics of the Cinnamyl Alcohol Dehydrogenase (CAD) and Cinnamate 4-Hydrolase (G4H) genes with basic wood density in *Neolamarckia cadamba*. Biotechnology, 11: 307-317. DOI: 10.3923/biotech.2012.307.317

Thompson, J.E. and S.C. Fry, 2001. Reconstructing of wall-bound xylogucan by transglycosylation in living plant cells. Plant J., 26: 23-34. DOI: 10.1046/j.1365-313x.2001.01005.x

Thumma, B.R., M.F. Nolan, R. Evans and G.F. Moran, 2005. Polymorphisms in Cinnamoyl CoA Reductase (*CCR*) are associated with variation in microfibril angle in *Eucalyptus* spp. Genetics, 171: 1257-1265. DOI: 10.1534/genetics.105.042028

Tiong, S.Y., S.F. Chew, W.S. Ho and S.L. Pang, 2014a. Genetic diversity of *Neolamarckia cadamba* using dominant DNA markers based on Inter-Simple Sequence Repeats (ISSRs) in Sarawak. Adv. Applied Sci. Res., 5: 458-463.

Tiong, S.Y., W.S. Ho, S.L. Pang and J. Ismail, 2014b. Nucleotide diversity and association genetics of xyloglucan endotransglycosylase/hydrolase (XTH) and Cellulose synthase (CesA) genes in *Neolamarckia cadamba*. J. Biol. Sci., 14: 267-375. DOI: 10.3923/jbs.2014.267.275

WAC, 2004. A tree species reference and selection guide: *Anthocephalus cadamba*. World Agroforestry Centre.

Wass, M.N., L.A. Kelley and M.J. Sternberg, 2010. 3DLigandSite: Predicting ligand-binding sites using similar structures. Nucleic Acids Res., 38: 469-473. DOI: 10.1093/nar/gkq406

White, A.R., Y. Xin and V. Pezeshk, 1993. Xyloglucan glycosyltransferases in golgi membranes from *Pisum sativum* (pea). Biochem. J., 294: 231-238.

Xu, W., P. Campbell, A.K. Vargheese and J. Braam, 1996. The *Arabidopsis* XET-related gene family: Environmental and hormonal regulation of expression. Plant J., 9: 879-889. DOI: 10.1046/j.1365-313X.1996.9060879.x

Ye, X., S.H. Yuan, H. Guo, F. Chen and G.A. Tuskan *et al*., 2012. Evolution and divergence in the coding and promoter regions of the populus gene family encoding xyloglucan endotransglycosylase/hydrolase. Tree Genet. Genomes, 8: 177-194. DOI: 10.1007/s11295-011-0431-1

Yokoyama, R. and K. Nishitani, 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. DOI: 10.1093/pcp/pce154

Yokoyama, R. and K. Nishitani, 2004. Genomic basis for cell-wall diversity in plants: A comparative approach to gene families in rice and *Arabidopsis*. Plant Cell Physiol., 24: 1025-1033. DOI: 10.1093/pcp/pch151

Yokoyama, R. J.K.C. Rose and K. Nishitani, 2004. A surprising diversity and abundance of xyloglucan endotransglycosylase/hydrolase in rice: Classification and expression analysis. Plant Physiol., 134: 1088-1099. DOI: 10.1104/pp.103

Zhong, R. and Z.H. Ye, 2009. Secondary cell walls. Encyclopedia Life Sci. DOI: 10.1002/9780470015902.a0021256