Theory and Practice of Science and Technology
2025, VOL. 6, NO. 1, 145-150
DOI: 10.47297/taposatWSP2633-456928.20250601

Genome-wide identification and expression analysis of the xyloglucan endotransglycosylase/hydrolase (XTH) gene family in castor bean

Yizhong Chen

College of Forestry, Southwest Forestry University, Kunming, 650224

ABSTRACT
The xyloglucan endotransglycosylase/hydrolase (XTH) family plays a key role in plant cell wall remodelling. However, XTH family genes have not been reported in castor. In this study, a total of 24 RcXTH genes were identified from castor, and all XTH proteins from castor and Arabidopsis were classified into classes I/II , IIIA, IIIB, and Ancetral on the basis of phylogenetic relationships. Gene structures and conserved motifs within each group were similar. All XTH proteins in this study contain a characteristic conserved structural domain (Glyco_hydro_16 and XET_C). Subcellular predictions indicate that XTH is located mainly in the cell wall and some in the cytoplasm, and analysis of the gene family amplification mechanisms suggests that genome-wide duplication and tandem duplication may be the main mechanisms leading to the amplification of the XTH gene family. Interestingly, the tandemly duplicated genes all belonged to group I/II, suggesting that tandem duplication is the main reason for the highest number of genes in this group. Based on the analysis of RNA-Seq data, the XTH genes showed different tissue-specific expression patterns. In addition, the XTH gene was differentially expressed in the duplicates in both CD and CH varieties, which was hypothesised to be the main reason for the phenotypic differences between the two castor varieties. We have conducted the first systematic analysis of the XTH gene family in castor, and the results of this study can be used as a theoretical basis for further research on the function of XTH genes in different castor varieties, in order to provide a theoretical basis for dwarf breeding research in castor.

KEYWORDS
Ricinus communis; Tissue expression; comparative genomics; XTH

1 Background

The plant cell wall serves as an external support structure for the cell, which regulates cell size and cell shape in response to plant growth and development. It consists mainly of polysaccharide polymers, including cellulose, hemifibres, pectin and hemicellulose[1]. Xyloglucan is the most important hemicellulose in the primary cell wall of dicotyledonous plants[2]. Hemicellulose plays many important roles in plants, such as participating in the shaping of plant form and structure, providing mechanical support for plants, and resisting biotic and abiotic attack. In recent years, a growing body of research has shown that cell wall loosening and rearrangement are critical events in processes such as cell proliferation, volume expansion, shape change and stress adaptation. In particular, hemicellulose and the remodelling of its structure are indispensable factors for plants to be able to cope with different environmental conditions.

Xyloglucan endotransglycosylase/hydrolase (XTH) is widely distributed in plant cells and is used for cellular remodelling by catalysing the cleavage and polymerisation of xyloglucan molecules to regulate cellular elasticity and ductility. Belongs to the glycoside hydrolase family 16 and is a key enzyme involved in remodelling plant cells[3]. The XTH family genes contain a typical catalase base (HDEIDFEFLG).

The first glutamate residue (E) is the affinity site and the second is the proton donor. XTHs can form a covalent glycosylase intermediate that is cleaved by water to perform a hydrolysis reaction, or they can enter the sugar substrate to perform transglycosylation. On the basis of structural features, the XTH family can be divided into four groups: group I/II, group IIIA, group IIIB and the ancestral group. The XTH studies reported so far indicate that XTH with glycosyltransferase activity is predominantly group I/II, whereas XTH with hydrolase activity is predominantly group IIIA. With the rapid development of plant genome sequencing technology, members of the XTH gene family have been identified and characterised in a wide range of plants. For example, Arabidopsis thaliana , rice , poplar , sweet potato , barley , and Camellia sinensis.

Castor oil plant (*Ricinus communis* L) was one of the world's most important woody oil crops. In our previous study, we identified many candidate intervals of key genes regulating castor height, including the XTH gene, but its characterisation remains unclear. In this study, Then we analysed the phylogenetic relationship, gene structure, chromosomal location, subgenomic distribution, and organisational expression pattern of protein sequences of XTH, which lays the foundation for further investigation of castor XTH gene function.

For a better understanding of this gene family in higher plants; these findings provide useful information for a better understanding of the function and evolution of this gene family in higher plants; these findings provide us with a new understanding of the function of RcXTH and lay the foundation for further genetic engineering and genetic improvement of castor bean.

2 Results

2.1 Genome-wide identification of Castor XTH

To identify the XTH gene in wild castor, we performed a BLAST search with the castor protein sequence using the amino acid sequence of Arabidopsis thaliana AtXTH8 as a query. Based on the XTH protein structural domains Glyco hydro 16 and XET_C (PF00722 and PF06955), 53 XTH homologues were identified from the castor protein sequence. From these, 24 XTH genes were identified by combining BLAST and HMMER results (redundant sequences and sequences lacking features of the XTH structural domain were excluded using the EMBL-EBL online tool).

In addition, the physicochemical properties of the proteins of 24 XTH genes were analysed. Among them, RcXTH4 encoded the largest amino acid length (970) and molecular weight (MW) of 108.50 kDa, while RcXTH16 encoded the smallest number of amino acids (158). The theoretical isoelectric points (PI) ranged from 4.89 to 9.28, while the instability coefficients ranged from 27.59 to 56.44.

In addition, only one protein had a total hydrophobicity mean (GRAVY) greater than 0 and was considered hydrophobic, while all other proteins had a total hydrophobicity mean less than 0 and were hydrophilic. The plant-mPLoc online tool was used to predict the subcellular location of RcXTH proteins. The results showed that 15 of the XTH proteins in castor were located in the cell wall. In addition to the cell wall, nine RcXTH proteins are located in the cytoplasm. Prediction of signal peptides of XTH proteins using the online tool TargetP-2.0 Server showed that 21 RcXTH have signal peptides (Table 1).

Table 1 Molecular characteristics of RcXTHs in grapevine

Gene name	Gene ID	AA	MW	PI	II	GRAVY	SignalP	Subcellular localization
RcXTH1	Rc01T001430.1	295	33790.66	5.61	45.31	-0.421	-	Cell wall
RcXTH2	Rc01T001890.1	277	31771.93	8.26	41.4	-0.365	S	Cell wall Cytoplasm.
RcXTH3	Rc01T002216.1	232	26174.22	6.88	30.86	0.002	S	Cell wall
RcXTH4	Rc03T005899.1	970	108504.56	8.86	36.03	-0.212	S	Cell wall
RcXTH5	Rc03T006470.1	399	45030.18	8.52	31.04	-0.24	S	Cell wall Cytoplasm.
RcXTH6	Rc03T006472.1	278	31346.22	7.58	40.77	-0.366	S	Cell wall
RcXTH7	Rc03T006473.1	284	31875.37	4.89	36.72	-0.422	S	Cell wall Cytoplasm.
RcXTH8	Rc03T006474.1	284	31764.27	5.22	35.86	-0.43	S	Cell wall Cytoplasm.
RcXTH9	Rc03T007052.1	290	33216.46	5.95	35.63	-0.324	S	Cell wall Cytoplasm.
RcXTH10	Rc05T009393.1	300	34841.06	5.23	32.31	-0.532	S	Cell wall
RcXTH11	Rc05T009925.1	350	40860.23	9.14	51.84	-0.565	S	Cell wall
RcXTH12	Rc05T011173.1	287	32805.28	8.68	41.64	-0.272	S	Cell wall Cytoplasm.
RcXTH13	Rc05T011174.1	294	33680.45	8.81	41.22	-0.262	S	Cell wall Cytoplasm.
RcXTH14	Rc05T011586.1	338	38362.23	7.08	45.94	-0.332	S	Cell wall
RcXTH15	Rc06T013591.1	289	33394.59	9.28	47.81	-0.598	S	Cell wall
RcXTH16	Rc06T013874.1	158	18387.02	8.67	56.44	-0.26	-	Cell wall
RcXTH17	Rc07T015114.1	293	33359.29	6.58	43.39	-0.587	S	Cell wall
RcXTH18	Rc07T015563.1	278	31446.44	8.24	27.59	-0.342	S	Cell wall Cytoplasm.
RcXTH19	Rc07T015565.1	293	32674.55	5.54	35.81	-0.251	S	Cell wall Cytoplasm.
RcXTH20	Rc09T019926.1	319	35780.32	6.86	54.24	-0.248	S	Cell wall
RcXTH21	Rc09T020151.1	842	94636.99	8.61	41.21	-0.282	S	Cell wall
RcXTH22	Rc09T021520.1	288	33193.57	6.5	44.58	-0.37	S	Cell wall
RcXTH23	Rc10T022203.1	362	41802.59	9.11	44.73	-0.247	-	Cell wall
RcXTH24	Rc10T024580.1	299	34646.18	8.45	36.03	-0.458	S	Cell wall

2.2 Phylogenetic analysis and classification of RcXTH

To analyse the evolutionary relationships between members of different XTH gene families, we used full-length XTH protein sequences from R. communis and A. thaliana to generate a phylogenetic tree based on the neighbour-joining (NJ) (Fig. 1). Four groups (Group I/II, Group IIIA, Grpup IIIB, and Ancetral) were identified based on the clade support values, the topology of the phylogenetic tree, and the previous classification of XTH families in Arabidopsis. Ancetral, which is close to the root, is the smallest group and contains six members. Group IIIA has 4 XTHs and Group IIIB has 9. The rest belong to group I/II and include 16 RcXTHs and 22 AtXTHs. there are 18 homologous pairs at the end of the branches of

the phylogenetic tree as shown in Figure 1.

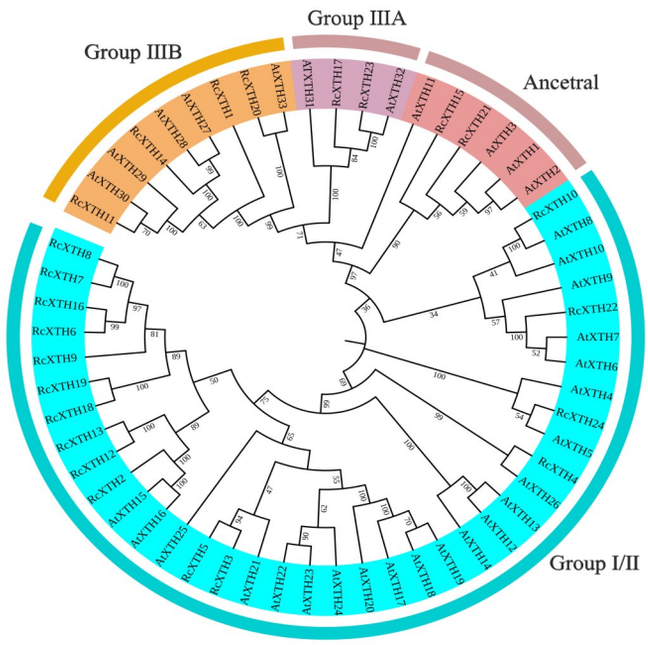


Figure 1 Phylogenetic analysis of castor and Arabidopsis XTH proteins, each colour represents a group

2.3 Structural, motif and protein structure analysis of the XTH gene in the XTH protein

Previous studies have shown that the exon organisation in the Arabidopsis XTH gene is well conserved. In order to better characterise the structural conservation and diversity of the XTH gene during evolution, exonic intronic organisation of the coding sequences of various groups of XTH genes was obtained from wild castor. Each XTH protein contains a Glyco_hydro_16 structural domain and an XET_C structural domain (as shown in Figures 2B and 2C, the Glyc_hydro_16 structural domain spans the 6-2-4-3-1 motif sequence, but some proteins lacked one or more protein motifs; the lengths of 2 XTHs (RcXTH16 and RcXTH3) were found to be less than 250 amino acids, due to the deletion of motifs 6, 2, 4, and 3 in the Glyc_hydro_16 structural domain (Fig. 2B)).The XET_C structural domain contains mainly motifs 5 and 8, with a similar distribution of conserved motifs in the same subgroup.

XTH All genes in Group I contain 3 or more introns except RcXTH16; all genes contain 2 to 3 exons except RcXTH3, which has no exons. All genes in Group II contain 4 or more introns and 2 exons. All genes in Group III contain 4 introns and 2 exons; all genes in Group 4 except RcXTH1 contain 4 introns and 2 exons. In general, the thematic patterns of the

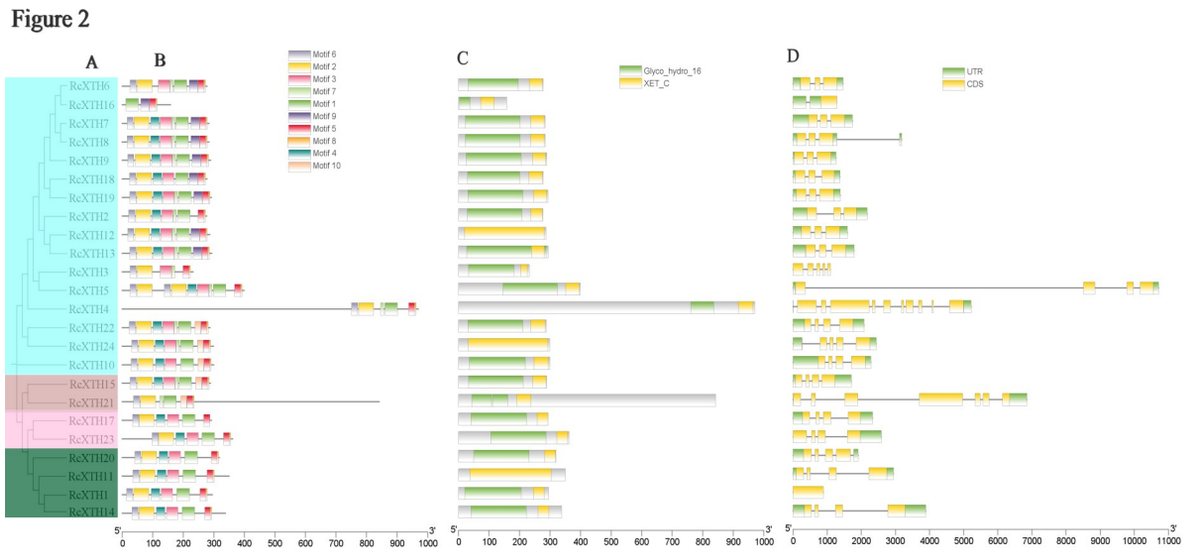


Figure 2 Conserved motifs, conserved protein structural domains and gene structure of the castor XTH gene. (A) Phylogenetic tree of the castor XTH gene family; (B) protein-conserved motifs; (C) conserved protein structural domains; (D) gene structure

different XTH proteins do not differ much.

2.4 Chromosomal localisation and duplication analysis of the XTH gene

The chromosomal locations of all the XTH genes in castor GFF3 file were analysed based on their physical locations, and the results are shown in Fig. 3, where 24 RcXTH genes were unevenly distributed on 10 chromosomes, with the highest number of genes (6) on Chr3 and the lowest number of genes (2) on Chr6 and Chr10. Interestingly, there is no RcXTH on chromosomes 2, 4 and 8.

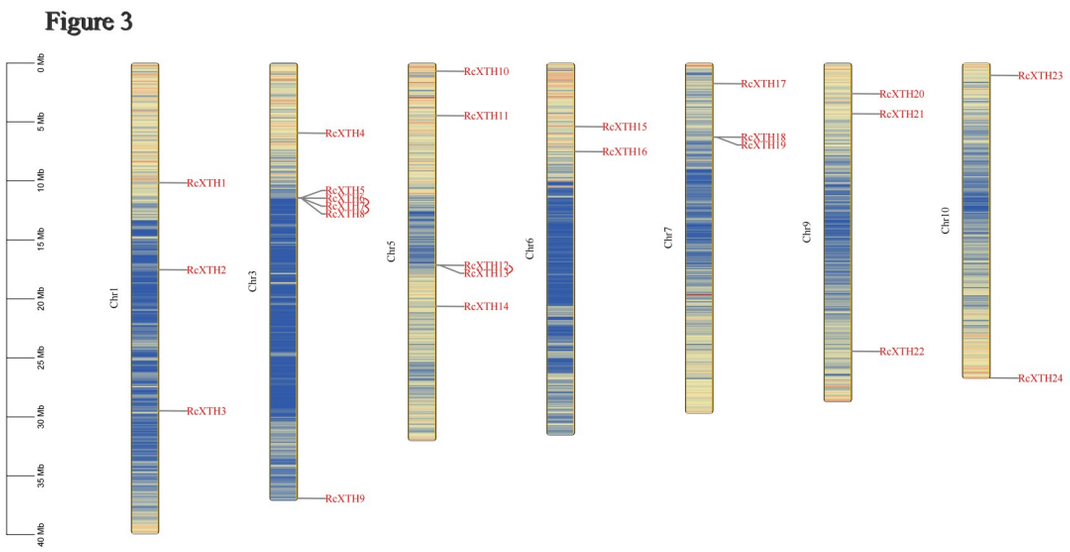


Figure 3 Locations of XTH genes on R. communis chromosomes

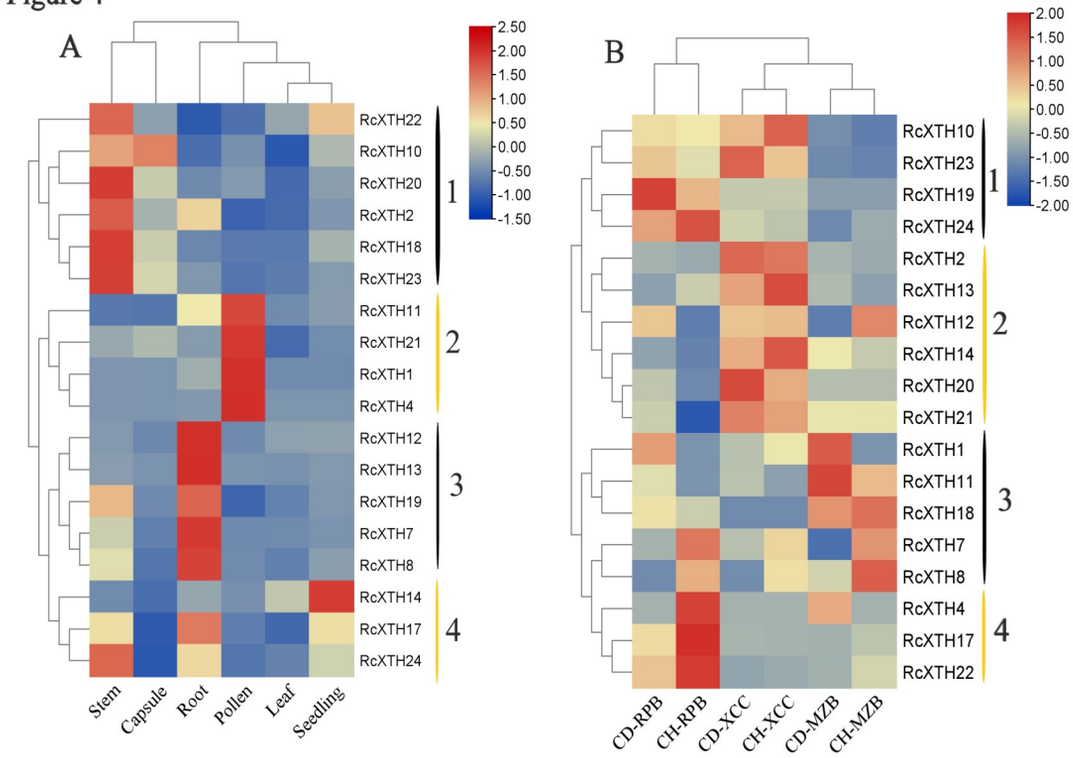
TD events contribute to the amplification of gene families and may generate tandemly duplicated genes in clusters. We inferred tandemly duplicated XTH genes by TBtools software and found three pairs of RcXTH genes in tandem arrays, which accounted for 25% of the total XTH genes in *R. communis*. These tandemly duplicated genes were clustered together, which is consistent with their location on the chromosome.

2.5 Expression pattern of XTH gene in different tissues of *R. communis*

In order to understand the variation in the expression pattern of XTH genes, we analysed the expression pattern of XTH genes in different tissues of two castor cultivars with significant differences in plant height based on RNA-Seq data . In this study, genes with FPKM values less than 1 were considered as unexpressed genes, including *RcXTH3*, *RcXTH5*, *RcXTH9*, *RcXTH15*, and *RcXTH16*. out of 24 RcXTH genes. On this basis, 20 XTH genes were expressed in at least one tissue, while the remaining genes were not expressed/lowly expressed in the remaining tissues, suggesting that they may be non-functional or specifically expressed in other tissues, but the present study indicated that. Among the 24 RcXTH genes, 8 XTH genes were expressed in all tissuesin all tissues. the remaining XTH genes were specifically expressed in at least one tissue, such as *RcXTH2*, *RcXTH22*, and RcXTH24 were specifically expressed in stems; *RcXTH19* was specifically expressed in roots; and RcXTH21 was specifically expressed in pollen (Fig. 4A).

Cluster analysis of expression values showed that the XTH genes of R. communis could be classified into four groups (Fig. 4). The XTH genes in group 1 had significantly higher expression in stems than in the other tissues; in group 2, the XTH genes were expressed in pollen than in the rest of the tissues; in group 3, the XTH genes were expressed in roots than in the rest of the tissues; similarly in group 4, RcXTH14 was specifically expressed in the seeds, and the expression of RcXTH24 in stems was higher than that of the rest of the genes in the group, and RcXTH17 was more highly expressed in roots (Fig. 4A). To explore the expression of XTH genes in stems, we used RNA-Seq of two castor varieties with significant plant height differences to analyse their expression patterns in stems. Based on the genes that were highly expressed in stems and roots (Fig. 4A), we found that RcXTH12 and RcXTH18 were more highly expressed in CH-MZB than CD-MZB; RcXTH20 and RcXTH23 were more highly expressed in CD-XCC than CH-XCC, whereas RcXTH13 and RcXTH14 were more highly expressed in CH-XCC than CD-XCC; RcXTH4, RcXTH17, RcXTH22 and RcXTH24 were higher expressed in CH-RPB than in CD-RPB; on the contrary, RcXTH19 was higher expressed in CD-RPB than in CH-RPB.It is hypothesised that the XTH genes may contribute to phenotypic differences in castor plants by being involved in the regulation of the associated growth of the roots and stems of castor.

Figure 4



3 Discussion

XTH family genes play a critical role in plant growth and development. In this study, we found 24 XTH genes in *R. communis*, which were sequentially named RcXTH1~RcXTH24 based on their positions on the chromosomes; the 24 RcXTH genes were classified into four groups, which were similar to those of Arabidopsis thaliana, rice and poplar. Each XTH gene contains two major conserved structural domains of XTH. In addition, we predict that most of the RcXTH proteins are located in the cell wall and a small fraction in the cytoplasm.

Plants have undergone many genome-wide duplication events during evolution, including tandem duplications, segmental duplications, and transformation events. In this study, we analysed the covariance events of castor XTH family genes, and we found that the number of tandemly duplicated gene pairs was less than the number of segmentally duplicated genes, suggesting that castor may have lost some of its XTH tandem genes during gene duplication. Tandem duplication favours the amplification of XTH gene families, such as wheat, tobacco and soybean.

In the present study, the XTH gene of *R. communis* was expressed in all tissues, and different XTH genes in the same subgroup showed different expression patterns, even for homologous genes with a high degree of amino acid sequence identity. *AtXTH21*, AtXTH22 and AtXTH30 were mainly expressed in fruits under the previous study hair. However, homologous genes in castor showed different expression patterns. For example, RcXTH7 was expressed in roots and stems, and RcXTH11 was expressed in pollen and roots. RcXTH20 had the highest expression in stems and CD-XCC, consistent with its homologue (*AtXTH33*). *RcXTH7*, as a homologue of *AtXTH24*, had the highest expression in roots, whereas the Arabidopsis homologue (*AtXTH24*) was expressed mainly in stems.

4 Conclusions

In this study, 24 XTH genes were identified based on the whole genome of castor. They were classified into four groups (I/II, IIIA, IIIB and Ancetral) with 33 XTH genes in Arabidopsis thaliana by phylogenetic analysis based on the clade support values, the topology of the phylogenetic tree, and the previous classification of Arabidopsis thaliana families. Comparison of exon intron distribution and conserved motifs also supported this classification of XTH. the WGD event had the greatest impact on the amplification of the number of genes in the XTH family in castor, followed by the tandem duplication event. the analysis of the RNA-Seq data provided insights into the functional specificity of the individual members of the XTH gene family across species. In conclusion, these results increase our understanding of the XTH gene family and provide a reference for dwarf breeding of castor in the future by functional characterisation of XTH genes that are highly expressed in the stem.

5 Materials and methods

5.1 Data sources

The genome sequence, CDS sequence, protein sequence and annotation information of R. communis were downloaded from the subject Euphorbiaceae DB database. The Arabidopsis XTH protein sequence was downloaded from TAIR.

5.2 Genome-wide identification of the XTH gene

In this study, we first obtained 33 Arabidopsis XTH sequences from the TAIR database as query sequences and performed a BLASTp (E < 1.0e-5) search for all protein sequences of R. communis from the Euphorbiaceae DB database.

5.3 Phylogenetic analysis of RcXTH proteins

Protein sequences of the Arabidopsis XTH family were downloaded from Phytozome13. ClustalW was used for multiple sequence comparison of proteins, and MEGA11 software was used to construct a phylogenetic tree by the neighbour-joining (NJ) method. Evolutionary trees were visualised by the itol online tool.

References

[1] ZHANG B, GAO Y, ZHANG L, et al. The plant cell wall: Biosynthesis, construction, and functions [J]. Journal of integrative plant biology, 2021, 63 (1): 251-72.

[2] COSGROVE D J, JARVIS M C. Comparative structure and biomechanics of plant primary and secondary cell walls[J]. Front Plant Sci, 2012, 3: 204.

[3] DONG J, MA Y, HU H, et al. The Function of SD1 on Shoot Length and its Pyramiding Effect on Shoot Length and Plant Height in Rice (Oryza sativa L.)[J]. Rice (New York, NY), 2024, 17(1): 21.