Identification and analysis of the expansin gene family in yam (#114649)

First submission

Guidance from your Editor

Please submit by **6 Apr 2025** for the benefit of the authors (and your token reward) .

Structure and Criteria
Please read the 'Structure and Criteria' page for guidance.

Raw data check
Review the raw data.

Image check
Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files

Download and review all files from the materials page.

11 Figure file(s)
2 Table file(s)
1 Raw data file(s)
9 Other file(s)

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- **Impact and novelty is not assessed.** Meaningful replication encouraged where rationale & benefit to literature is clearly stated.
- All underlying data have been provided; they are robust, statistically sound, & controlled.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Identification and analysis of the expansin gene family in yam

Xing Qiao [1, 2, 3], **Shuai Yuan** [1, 2, 3], **Jing Wei** [1, 2, 3], **Chang zhou Li** [1, 2, 3], **Lan Lan** [1, 2, 3], **Xue rui He** [1, 2, 3], **Ai Qin Wang** [1, 2, 3], **Jie Zhan** [1, 2, 3], **Long fei He** [Corresp., 1, 2, 3], **Dong Xiao** [Corresp. 1, 2, 3]

[1] College of Agriculture, Guangxi University, Nanning, Guangxi Zhuang Autonomous Region, China

[2] National Demonstration Center for Experimental Plant Science Education, Nanning, China

[3] Guangxi University Key Laboratory of Crop Cultivation and Tillage, Nanning, China

Corresponding Authors: Long fei He, Dong Xiao
Email address: lfhe@gxu.edu.cn, xiaodong@gxu.edu.cn

Expansins are a group of proteins that loosen plant cell walls and cellulose materials and are involved in regulating plant cell growth and diverse developmental processes. However, a systematic study of the Dioscorea opposita expansin (DoEXP) gene family has not yet been conducted. In this study, we used publicly available genomic data from yam to characterize the DoEXP gene family and analyse its physicochemical properties, phylogeny and expression pattern using bioinformatics software. Thirty EXP genes were identified from the yam genome and can be classified into four subfamilies, DoEXPA, DoEXPB, DoEXLA, and DoEXLB, which are distributed across 14 chromosomes, and all EXP proteins contain two conserved structural domains: DPBB_1 and expansin_C. Structural analyses of the DoEXP gene family revealed that they have similar motif compositions and exon–intron structures. A cis-acting progenitor analysis revealed that DoEXP family members contain response elements for plant growth and development, the light response, phytohormones, and low-temperature and drought stresses, suggesting that they may play roles in plant growth and development and abiotic stress responses. The collinearity analysis revealed multiple pairs of collinear genes within the DoEXPA subfamily, indicating that the gene evolution of this subfamily occurred through the replication of chromosomal segments. Moreover, an analysis of the evolutionary pressure of EXP genes within and between species revealed that the genes in this family were subject to purifying selection. The functional enrichment analysis of the DoEXP genes revealed that the yam expansin genes play significant roles in root elongation and cell differentiation. The reliability of the RNA-seq data was confirmed by qPCR results, which further validated the expression patterns of DoEXP genes. The results of this study are helpful for understanding the molecular functions of expansin proteins in yam tuber expansion and provide a theoretical basis for revealing the molecular regulatory mechanism of yam tuber growth and development.

Identification and Analysis of the Expansin Gene Family in Yam

Xing Qiao[1,2,3], Shuai Yuan[1,2,3], Jing Wei[1,2,3], Changzhou Li [1,2,3], Lan Lan[1,2,3], Xuerui He[1,2,3], Aiqin Wang[1,2,3], Jie Zhan[1,2,3],Longfei He [1,2,3] *, Dong Xiao[1,2,3] *

[1]Guangxi Key Laboratory of Agro-environment and Agro-products Safety, College of Agriculture, Guangxi University, Nanning, Guangxi, 530004, China
[2] National Demonstration Center for Experimental Plant Science Education, Nanning 530004, China
[3] Guangxi University Key Laboratory of Crop Cultivation and Tillage, Nanning 530004, China

Corresponding Author:
Longfei He[1,2,3] *, Dong Xiao[1,2,3] *
Guangxi University, Nanning, Guangxi, 530004, China
Email address: Dong Xiao: xiaodong@gxu.edu.cn, Longfei-He: lfhe@gxu.edu.cn.

Abstract:

Expansins are a group of proteins that loosen plant cell walls and cellulose materials and are involved in regulating plant cell growth and diverse developmental processes. However, a systematic study of the *Dioscorea opposita* expansin (DoEXP) gene family has not yet been conducted. In this study, we used publicly available genomic data from yam to characterize the DoEXP gene family and analyse its physicochemical properties, phylogeny and expression pattern using bioinformatics software. Thirty EXP genes were identified from the yam genome and can be classified into four subfamilies, DoEXPA, DoEXPB, DoEXLA, and DoEXLB, which are distributed across 14 chromosomes, and all EXP proteins contain two conserved structural domains: DPBB_1 and expansin_C. Structural analyses of the DoEXP gene family revealed that they have similar motif compositions and exon–intron structures. A cis-acting progenitor analysis revealed that DoEXP family members contain response elements for plant growth and development, the light response, phytohormones, and low-temperature and drought stresses, suggesting that they may play roles in plant growth and development and abiotic stress responses. The collinearity analysis revealed multiple pairs of collinear genes within the DoEXPA subfamily, indicating that the gene evolution of this subfamily occurred through the replication of chromosomal segments. Moreover, an analysis of the evolutionary pressure of EXP genes

within and between species revealed that the genes in this family were subject to purifying selection. The functional enrichment analysis of the DoEXP genes revealed that the yam expansin genes play significant roles in root elongation and cell differentiation. The reliability of the RNA-seq data was confirmed by qPCR results, which further validated the expression patterns of DoEXP genes. The results of this study are helpful for understanding the molecular functions of expansin proteins in yam tuber expansion and provide a theoretical basis for revealing the molecular regulatory mechanism of yam tuber growth and development.

Keywords: Yam, Expansin family, Evolutionary analysis, Gene function, Expression patterns, Tuber expansion

Background

Expansins are a group of proteins that loosen plant cell walls and cellulose materials and are involved in regulating plant cell growth and diverse developmental processes. The EXP gene families have been extensively studied in many plant species. Although the whole genome sequence of the yam (*Dioscorea opposita*) is available, a comprehensive analysis of its EXP gene family remains lacking.

Methods

Using bioinformatics tools, expansin (EXP) genes were identified from the whole genome sequence of yam. A comprehensive analysis of the DoEXP gene family was conducted, including their physicochemical properties, phylogenetic relationships, gene structures, conserved motifs and domains, cis-acting element analysis, and collinearity analysis. The expression patterns of DoEXP genes at different growth and developmental stages were analyzed using RNA-seq data obtained in the laboratory, and the results were validated by real-time quantitative polymerase chain reaction (qPCR).

Results

In this study, 30 DoEXP genes were identified and classified into four subfamilies, which were distributed across 14 chromosomes. All these genes contained two conserved domains (DPBB_1 and expansin_C). Genes within the same subfamily exhibited similar motif compositions and exon-intron structures. Cis-acting element analysis suggested that these genes play roles in plant growth and development, abiotic stress responses, and hormonal responses.

Collinearity analysis revealed the presence of gene duplication events, particularly in the DoEXPA subfamily, while evolutionary analysis indicated that the expansin gene family was under purifying selection. Functional enrichment analysis highlighted the important roles of these genes in root elongation and cell differentiation, providing insights into the molecular mechanisms underlying the tuber expansion of yam. This study lays a foundation for further functional research and breeding of high-yield yam varieties.

Introduction

Yam (*Dioscorea opposita*) is a monocotyledonous plant belonging to the genus *Dioscorea* in the family *Dioscoreaceae*. It is harvested from underground tubers and ranks among the world's top ten food crops. Additionally, it is an important medicinal and food crop in China and is valued for its economic importance (Zhou et al., 2021). Its tubers are rich in starch, cellulose, amino acids, sugars, fats, and other nutrients, as well as medicinal components such as allantoin, diosgenin, polysaccharides, and polyphenols (Obidiegwu et al., 2020). These components tonify the spleen, nourish the stomach, strengthen the lungs, and benefit the kidneys. The tubers can be used to treat conditions such as prolonged diarrhoea due to deficient spleen function, chronic intestinal inflammation, gastritis, and diabetes mellitus. The growth and development of tubers affect the yield and quality of yam, and the function of expansins is closely related to the expansion of yam tubers.

Expansins (EXPs) are cell wall-associated proteins that participate in cell wall loosening and cell enlargement in a pH-dependent manner (Cosgrove, 2000). Research has confirmed that expansins are involved in plant growth and abiotic stress responses (Jin et al., 2020; Choi et al., 2003). The first expansin was discovered in cucumber (Shcherban et al., 1995). Expansins are ubiquitously found in various plants (Wu et al., 2001; Lin et al., 2005; Shi et al., 2014; Gao et al., 2020). In plants, expansins are involved in cell wall relaxation. The growing cell wall is structured as a multilayered arrangement of nearly parallel cellulose microfibrils tethered by hemicellulose. Expansins mediate the disruption of these microfibrils, leading to cell wall relaxation and facilitating cell elongation (Marga et al., 2005). The expansin superfamily can be divided into four families: α-expansins (EXPA), β-expansins (EXPB), expansin-like A (EXLA), and expansin-like B (EXLB) (Kende et al., 2004). Typically, expansins consist of 250–275 amino acids, including two conserved domains, DPBB_1 and expansin_C, and a signal peptide (SP) region located at the N-terminus with a length of 20–30 amino acids. Domain 1 is a

conserved region featuring a double-psi β-barrel (DPBB) fold located in the catalytic domain in the middle of the entire protein sequence. It contains 120–135 amino acids, is rich in cysteine residues, and shares high homology with glycoside hydrolase family 45 (GH45). The expansin_C domain is a stable domain located at the carboxyl terminus of expansins and generally contains 90–120 amino acids. This domain has a certain degree of homology with Group 2 grass pollen allergen proteins (G2A) in *Poaceae* (Sampedro & Cosgrove, 2005).

Expansins play essential roles in plant growth and development. Studies have shown that *OsEXPA10*, an Al-induced expansin gene, is involved in rice root cell elongation (Che et al., 2016); the overexpression of *AcEXPA23* promotes lateral root development in kiwi fruit (Wu et al., 2022). Overexpression of the winter wheat extensin gene *TaEXPA7-B* in rice results in salt tolerance and significantly promotes the growth of lateral root primordia and cortical cells (Wang et al., 2024). *HvEXPB7* enhances drought tolerance by increasing root hair growth in barley under drought stress (He et al., 2015). Overexpression of the wheat expansin gene *TaEXPA2* improves seed production and drought tolerance in transgenic tobacco plants (Chen et al., 2016). The Arabidopsis expansin gene *AtEXPA18* can improve drought tolerance in transgenic tobacco plants (Abbasi et al., 2021). Auxin-mediated *CqEXPA50* expression promotes salt tolerance in quinoa (*Chenopodium quinoa*) through interactions with auxin pathway genes (Sun et al., 2022). Overexpression of the wild soybean expansin gene *GsEXLB14* increases the tolerance of transgenic soybean hairy roots to salt and drought stress (Wang et al., 2024). The overexpression of *AcEXPA1* enhances aluminium tolerance in carpetgrass through the regulation of root growth (Li et al., 2024). These studies increase our understanding of expansin functions, revealing their essential roles not only in normal plant growth and development but also in the response to abiotic stresses (such as salt, drought, and aluminium toxicity) and plant hormones (such as auxin). However, for *Dioscorea opposita* (yam), the knowledge of expansins, including the characteristics of family members, expression patterns, and functional features, remains limited. This study identified and analysed EXP family members in yam using the yam genome database, which provides a basic basis for future analyses of whether the EXP gene family regulate the growth and developmental processes and stress tolerance pathways in yam.

Materials & Methods

Identification of Yam EXP Family Members, Analysis of Physicochemical Properties,

and Construction of Phylogenetic Trees

The Dalata_550_v2.1 file was downloaded from the Phytozome database (https://phytozome-next.jgi.doe.gov/) to obtain the yam genome file and protein sequences and identify the expansin (EXP) genes in yam. We employed two different methods to identify the members of the EXP gene family in yam. First, by retrieving the EXP genes from the Arabidopsis database(https://www.arabidopsis.org/), the sequences of 35 EXP genes were obtained and translated into amino acid sequences to be used as query sequences. A sequence alignment and screen were conducted on the whole-genome database of yam using TBtools, with an E value of $1e^{-5}$ used as the screening threshold. The amino acid sequences of the Arabidopsis EXP genes were subsequently used to retrieve conserved domains through PFAM(http://pfam.xfam.org/). The genes of the expansin family all contain two conserved domains, namely, DPBB_1 (PF03330) and expansin_C (PF01357). Next, we used hidden Markov models (HMMs) with PFAM numbers PF03330 and PF01357 to perform Hmmsearch on the yam protein dataset, with an E value $\leq 10^{-5}$. The candidate genes obtained using the two methods were combined, and the overlapping genes were screened. The protein sequences of the expansin family genes in yam were subsequently obtained. The molecular weights (MWs), isoelectric points (pIs), and hydrophilicity of each protein sequence were obtained from ExPaSy(https://web.expasy.org/compute_pi/), and their subcellular localizations were predicted using the online software Plant-mPLoc(http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/).

The protein sequences of the EXP genes from four species, namely, yam, *Arabidopsis thaliana*, wheat, and rice, were subjected to multiple sequence alignment using the ClustalW function of MEGA 11 with the default parameters, and a neighbour-joining (NJ) phylogenetic tree with 1000 bootstrap replicates was constructed. The phylogenetic tree was subsequently constructed using the iTOL website(https://itol.embl.de/itol.cgi).

Analysis of Conserved Protein Motifs, Stable Domains, and Gene Structures of Yam EXP Family Members

MEME online software(https://meme-suite.org/meme/tools/meme) was used to identify the conserved motifs existing in the 30 EXP genes of yam with the default parameter settings. The number of motifs was set to 10. After the output file was obtained, TB tools were used to visualize the conserved protein motifs retrieved from the MEME website. The stable domains and their positions contained in the yam EXP genes were retrieved from the Interpro database(https://www.ebi.ac.uk/interpro/). Domains with PFAM numbers recorded in the

database were used as screening conditions, and the domain data contained in the EXP genes were manually sorted. TBtools was subsequently used to visualize the data. Diagrams of the exon–intron structure were drawn based on the annotation information of the yam genome and the candidate gene IDs.

Chromosomal localization of the genes and analysis of cis-acting elements in the promoters of *DoEXPs*

The information of the yam genome annotation file was used to analyse the chromosomal localization of renamed yam EXP genes with TB tools. The bin size was set to 100 kb, and the gene density information file for each chromosome was output and displayed in the chromosome frames in the figure.

The gene sequence 2000 bp upstream of the target gene was extracted by TB tools software, and the genomic data of yam and the response elements present in the sequence were analysed using the online software Plantcare(http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The analysis was performed to extract cis-acting elements in the promoters that are related to growth and development, phytohormones, and the stress response using the output of the website, and the data were subsequently visualized using TB tools. The number of major cis-acting elements contained in each EXP gene was counted, and a heatmap was drawn.

Covariance analysis of EXP genes

Intraspecific Collinearity Analysis: The McscanX function of TBtools was used to study the tandem and segmental duplications of the yam EXP family genes. The bin size was set to 100 kb to obtain the gene densities on chromosomes. With each occurrence of a gene set to a value of 1, the number of gene occurrences was summed. Heatmaps and line graphs were drawn to represent the gene density on each chromosome. Finally, the collinear relationships among the yam EXP genes were visualized using the Advanced Circos function. Nucleotide sequences with an alignment identity and similarity rate greater than 75% and a distance between genes on the same chromosome less than 100 kb were selected as tandem duplications. Genes located in the duplicated regions and in the nucleotide sequences with alignment rates greater than 75% were selected as the result of segmental duplications (Cannon et al., 2004). The gene pairs derived from tandem duplication and segmental duplication that were obtained from the analysis were subjected to an analysis of selection pressure using the TBtools Simple Ka/Ks Calculator (NG). The ratio of the nonsynonymous substitution rate (Ka) to the synonymous substitution rate (Ks) was calculated. If Ka >> Ks or Ka/Ks >> 1, the gene was under positive selection; if Ka = Ks or

Ka/Ks = 1, the gene underwent neutral evolution; if Ka << Ks or Ka/Ks << 1, the gene was under purifying selection. The Ks value was used to estimate the approximate date of each duplication event that occurred in yam with the following formula: $T = Ks/2\lambda \times 10^{-6}$ Mya (million years ago) ($\lambda = 6.5 \times 1^{-9}$) (Chang et al., 2023).

Interspecific Collinearity Ana0lysis: The collinearity among three species, namely, the monocotyledonous plant rice, the dicotyledonous plant *Arabidopsis thaliana*, and the monocotyledonous plant yam, was analysed. The genome and gene annotation files of rice were sourced from the Ensembl Plants online website(https://plants.ensembl.org/index.html). The source of the download of the genome and annotation files of yam was introduced earlier in the article and will not be repeated here. The McscanX function of TBtools was used to conduct interspecific collinearity analyses between rice, *A. thaliana* and yam, with the default parameters. The DoEXP genes that had collinear relationships with those of the other two species were highlighted for visualization. Moreover, the collinear gene pairs between yam and rice and between yam and *A. thaliana* were identified, the evolutionary selection pressure values (Ka/Ks) were calculated, and scatter plots were drawn using ggplot2.

Heatmap and analysis of the expression of *DoEXPs*

The transcriptome data of the existing varieties in the laboratory were used as the source data to analyse the expression levels of *DoEXP* genes. These data were obtained by measuring the gene expression levels of the yam variety NH1 (abbreviated as NH) in four growth stages, and each stage was assessed as three technical replicates to ensure the relative accuracy of the data. The gene expression levels are presented as the FPKM values. Next, the expression levels of the target genes in this study, which were measured three times in each of the four periods, were sorted and summarized. Since the results of the transcriptome data revealed that two genes, *DoEXPA14* and *DoEXPA15*, were not expressed, they were temporarily not shown in the subsequent plot. Finally, the HeatMap function of TBtools was used to visualize the data. Using its built-in row normalization and Euclidean clustering methods, the trends for the changes in the expression levels of each gene in different periods were displayed, and the genes with similar expression patterns were clustered.

GO Functional Enrichment Analysis of *DoEXPs*

The yam genome was annotated at the whole-genome level using the online website EggNOG-mapper(http://eggnog-mapper.emblde/). A Gene Ontology (GO) enrichment analysis was conducted using the R package clusterProfiler. Pathways with $P < 0.05$ were selected as

significantly enriched pathways, and bubble plots were drawn using ggplot2.

Plant matrials

The experimental material used in this study was the "NH1" variety, cultivated at the Agricultural Demonstration Base of Guangxi University. Planting was conducted at the end of April 2024. Samples were collected at the tuber formation stage (NH-F) four months after planting. Subsequent sampling was performed monthly, covering the following stages: early tuber swelling (NH-E), middle tuber swelling (NH-M), and late tuber swelling (NH-L). Random sampling was employed to ensure representativeness. At each growth stage, three tuber samples were collected. These samples were sliced thinly using a sterile knife, and slices from the three tubers were pooled to form one experimental sample, Three biological replicates were prepared for each stage. The samples were wrapped in aluminum foil, labeled clearly, and immediately flash-frozen in liquid nitrogen. They were then stored at -80°C in an ultra-low temperature freezer for subsequent analysis.

RNA extraction, cDNA synthesis, qRT-PCR, and Expression analysis

The Vazyme FastPure Universal Plant Total RNA Isolation Kit (Vazyme, Nanjing, China) was used for RNA extraction. RNA integrity was confirmed by 1.2% agarose gel electrophoresis. RNA concentration and purity were measured using a NanoDrop ONE spectrophotometer (Thermo Scientific, USA). First-strand cDNA was synthesized from 1.0 μg of total RNA using the HiScript III All-in-one RT SuperMix Perfect for qPCR Kit(Vazyme, Nanjing, China). The resulting cDNA was diluted three-fold and stored at −20 °C for subsequent experiments for subsequent quantitative PCR (qPCR) experiments. Specific primers for the DoEXP genes were designed using Primer 5 software(see Table S4), and their specificity was verified using NCBI Primer-BLAST. Using the actin gene *DoActin* (GenBank accession no. KU669295) as the reference gene. Each qPCR reaction was performed in a 10 μL volume containing 1.0 μL of cDNA, 0.4 μL each of forward and reverse primers (10 μM final concentration), 3.2 μL of nuclease-free water, and 5.0 μL of AceQ qPCR SYBR Green Master Mix (Vazyme, Nanjing, China). Reactions were carried out on a BIO-RAD CFX96™ Real-Time System with the following program: 95°C for 3 minutes , followed by 39 cycles of 95°C for 10 seconds and 55°C for 30 seconds with plate reading. A melt curve analysis was performed from 65°C to 95°C with a 0.5°C increment and a 5-second hold at each step. The expression levels were calculated by the comparative Ct method, as reported by Schmittgen (Livak & Schmittgen, 2001),with the expression levels of each gene during the NH-F stage serving as the control group, and other

stages as the experimental group for analysis. At least three technical replicates and three independent biological replicates were performed for each sample. The data were analysed with one-way ANOVA using SPSS Statistics (version 22.0, IBM Corporation), with $P<0.05$ indicating significant differences. The results were plotted using GraphPad Prism 8.0.

Results and Analysis

Identification and evolutionary analysis of expansin gene family members in yam

In this study, 30 members of the expansin gene family were screened from the whole genome of yam. All expansin genes were classified into four subfamilies and named sequentially according to their chromosomal locations. Similar to Arabidopsis, rice, wheat, and other plants, the EXPA subfamily constituted the largest branch, with 21 members, followed by 4 EXLA genes, 3 EXLB genes, and 2 EXPB genes (Table 1). The start and end positions of the identified yam expansin genes and the positive and negative strands of the genes are shown. The relative molecular weights (MWs) of the proteins encoded by the full-length EXP genes ranged from 22016.99 (*DoEXPA12*) to 31599.85 (*DoEXPA2*) Da, and the lengths ranged from 206 (*DoEXLA12*) to 289 (*DoEXPA2*) amino acids (aas). Additionally, the predicted isoelectric points (pIs) of the expansins ranged from 4.43 (*DoEXLB1*) to 9.66 (*DoEXPA4*). Among them, five expansins have theoretical pIs less than 7, classifying them as acidic proteins, whereas the remaining 25 expansins are alkaline proteins. The hydrophilicity values were calculated based on the amino acids of the proteins. Nineteen expansin proteins had negative hydrophilicity values, indicating that they were hydrophilic proteins, among which *DoEXPA3* was the most hydrophilic. Eleven expansin proteins had positive hydrophilicity values, indicating that they were hydrophobic proteins, and *DoEXPA13* was the most hydrophobic. The analysis of subcellular localization revealed that all of the expansin proteins were located in the cell wall, a result that corresponds to the definition of expansin proteins as a group of proteins that loosen the cell wall and cellulose material of plants. Detailed information on the expansin genes in yam (*DoEXPs*) is provided in Table 1.

This study selected expansin family genes from four species to analyse the evolutionary relationships of the expansin gene families among different species: *Dioscorea opposite* (yam), *Arabidopsis thaliana*, *Triticum aestivum L.* (wheat), and *Oryza sativa L.* (rice). Using the TAIR online website, the expansins in *A. thaliana* were identified, and data on the reported expansins

in *O. sativa* and *T. aestivum* were collected from previously published articles. We finally constructed a phylogenetic tree using 119 expansin proteins (including 34 in Arabidopsis, 30 in yam, 37 in rice, and 18 in wheat) collected from the four species (Fig. 1). The expansin genes of different species were clearly divided into four subfamilies, EXPA, EXPB, EXLA, and EXLB, and the members of each of the four subfamilies clustered together.

Conserved protein motifs, stable structural domains and analysis of the *DoEXP* gene structures

The online software MEME was used to analyse the amino acid sequences of DoEXPs for conserved protein motifs, and the number of motifs was set to 10. The visualization of the results (Fig. 2) revealed that members of the same subfamily were similar in terms of the type, number and order of conserved motifs, but some divergence was observed between the different subfamilies, which indicated that the yam expansin genes have diverse biological functions but also share certain connections that reflect the conservation of this family during evolution. The EXPA subfamily is the largest subfamily and has a more stable motif composition, except for *DoEXPA11, DoEXPA12, DoEXPA14, DoEXPA19, and DoEXPA20*, which are highly conserved, and we speculated that these genes may have similar functions. The EXPB subfamily shares the same motif structure: motif1, motif4, motif5, motif6, motif7, and motif8. The subfamilies EXLA and EXLB have the unique motif9, suggesting that these two subfamilies might be closer relatives than the subfamilies EXPA and EXPB are. A unified analysis of the conserved motifs of the four subfamilies revealed that all EXP genes possessed two conserved motifs, motif 1 and motif 6, indicating that these two motifs remained highly conserved during the evolution of expansin genes. The locations of the two stable domains in the yam expansin genes were in the middle and carboxyl-terminal (C-terminal) regions of the protein sequences.

The structure of introns and exons, regarded as the skeletal framework of genes, plays a significant role in regulating gene expression. The results (Fig. 2) showed that most members of the same subfamily presented similar gene structure compositions. The genes in the EXPA subfamily contained 2–3 exons, with only three genes (10%) containing two exons, and the majority of the genes (60%) containing three exons. Additionally, the number of introns in the EXPA subfamily ranged from 12, whereas the number of introns in the other three subfamilies ranged from 34, with the number of exons ranging from 45. Specifically, all EXPB subfamily genes contained four exons, and among the seven genes in EXLA and EXLB, six genes (20%)

contained five exon structures. These observations suggest that EXLA and EXLB might be phylogenetically closer. The above results indicate that the members within each subfamily share a high degree of structural conservation.

Chromosomal localization of *DoEXPs* and analysis of cis-acting elements

The locations of the DoEXP genes on the chromosomes were analysed and visualized using TBtools software (Fig. 3). The results indicated that the EXP genes were widely distributed on chromosomes, with members of EXP genes detected on all 14 chromosomes, among which chromosome 19 had the greatest number of EXP genes, with five, including three EXPA genes, one EXPB gene and one EXLB gene. Chromosome 14 followed with 4 EXP genes, while the number of EXP genes on other chromosomes ranged from 1 to 3. From the gene density heatmap of chromosomes, we concluded that the genes were mostly distributed at both ends of chromosomes, and the gene distribution was the lowest in the middle position of each chromosome. *DoEXLA1-DoEXLA2* on chromosome 2 and *DoEXPA14-DoEXPA15* on chromosome 15 are tandemly duplicated gene pairs.

We analysed the cis-acting elements present in the 2000 bp upstream promoter regions of 30 putative EXP genes to further understand the potential functions and regulatory mechanisms of the DoEXP family genes . A total of 13 major cis-acting elements were detected in this study, including light-responsive elements; response elements related to plant hormones such as gibberellin, salicylic acid, and auxin; and response elements related to abiotic stress responses. As shown in Figure 4, among the 30 yam EXP genes, 20 genes contained light-responsive elements, and the number of light-responsive elements was the greatest. Second, the EXP genes also contained many plant hormone-responsive elements, including 38 methyl jasmonate (MeJA) response elements. The responsiveness to MeJA is reflected mainly in the ability of plants to stimulate defence mechanisms and help plants resist abiotic stress environments. Three genes, *DoEXPA3, DoEXPA6*, and *DoEXPA18*, each have four MeJA response elements. Thirty-seven abscisic acid (ABA) response elements were identified, among which *DoEXPA18* contains as many as four ABA response elements, and *DoEXPA12* and *DoEXPA13* have three abscisic acid response elements. We speculated that *DoEXPA18* is strongly induced by ABA and MeJA signals. Thirty-one gibberellin (GA) response elements were identified and play important roles in plant cell division, root growth, and abiotic stress. Sixteen salicylic acid (SA) response elements were identified. As one of the nine major plant hormones, salicylic acid plays an

important role in the responses of plants to biological and abiotic stresses, such as disease resistance and stress resistance, and in the growth and development of roots. Twelve auxin (IAA) response elements were also identified.

In erms of abiotic stress, 12 defence and stress response elements and 20 drought-inducible (MBS) and 14 low-temperature-responsive (LTR) MYB binding sites were identified. Seven genes, *DoEXPA7, DoEXPA9, DoEXPA13, DoEXPA17, DoEXPB1*, *DoEXLA2*, and *DoEXLB1*, simultaneously contain these two abiotic stress response elements. These genes may play important roles in the abiotic stress response of yam. In addition to the above cis-acting elements, 8 circadian rhythm control elements, 16 meristem expression elements, 7 endosperm expression elements, and 2 cell cycle regulation elements were also detected, indicating that *DoEXPA1* and *DoEXPA8* have specificity in the response to cell cycle regulation. The functional elements identified above are functionally concentrated and affect plant morphology.

Intraspecific and Interspecific Collinearity of *DoEXPs*

Segmental duplication and tandem duplication are the main reasons for the expansion of plant gene families. We detected the collinearity among the yam EXP genes, between yam and rice, and between yam and Arabidopsis to investigate the evolutionary relationships of the EXP gene family. The results of the intraspecific collinear relationships are shown in Fig. 6, where the grey lines represent all covarying gene pairs occurring in the yam genome. For the four subfamilies of yam EXP genes, covarying gene pairs were observed only within the EXPA subfamily, and no covariance relationships were detected between genes in the other three subfamilies. On chromosome 7, the number of collinearities of the DoEXPA genes was the largest, with 8 collinear relationships. Both chromosomes 18 and 19 had six colinear gene pairs; five colinear relationships existed on chromosome 9, followed by three, three, and two colinear gene pairs on chromosomes 5, 12, and 14, respectively. A total of 2 pairs of tandemly duplicated gene pairs and 16 pairs of segmentally duplicated gene pairs were detected for DoEXPs, and the gene duplication information is listed in Table S1. These results indicate that the evolution of the DoEXPA genes occurred through large-scale chromosomal segmental duplications.

A total of 18 pairs of tandem duplications and segmental duplications were detected in yam. The ratio of the nonsynonymous substitution rate (Ka) to the synonymous substitution rate (Ks) was calculated. As shown in Table 2, except for the Ka/Ks values of the segmental duplication gene pairs DoEXPA1–DoEXPA4 that could not be calculated, the Ka/Ks values of the remaining

17 pairs of duplication gene pairs were all less than 1, indicating that the DoEXP gene family underwent purifying selection during the evolutionary process. Moreover, the divergence time of the duplication events was calculated using a formula to be approximately 3.9–17.8 million years ago (Mya).

The results of the interspecific evolutionary relationships are shown in Figure 7. A comparison of collinearity was conducted among species with the dicotyledonous plant Arabidopsis and the monocotyledonous plant rice. The results revealed that both Arabidopsis and rice have genes homologous to DoEXPs. Twenty-six pairs of orthologous gene pairs were identified between rice and yam. Although these 26 homologous gene pairs are collinear, the positions of the genes on the chromosomes differ slightly between yam and rice. Thirty-nine pairs of homologous genes were identified between yam and Arabidopsis, which was more than the number of homologous gene pairs between yam and rice. The Ka, Ks, and Ka/Ks values of the homologous gene pairs between yam and rice and between yam and Arabidopsis were subsequently calculated (Tables S2 and S3). In Arabidopsis and yam, the Ka/Ks values of 11 pairs of homologous genes could not be calculated, and the Ka/Ks values of the remaining 28 pairs of homologous genes were all less than 1.0. An analysis of the homologous gene pairs between yam and rice revealed that the Ka/Ks values of 16 pairs could not be calculated, and the Ka/Ks values of the remaining homologous gene pairs were all less than 1.0. The above results indicate that expansin gene family has undergone strong purifying selection during the evolutionary process. The Ka/Ks values of all collinear gene pairs within and between species were plotted (Fig. 8).

Analysis of *DoEXP* expression

Using transcriptome data, the expression levels of EXP genes at different time points were visualized in a heatmap. Most notably, the expression pattern of the *DoEXPA19* gene at the four different periods completely differed from that of the other genes, and the expression of this gene at the late stage of expansion was the highest among the four periods; thus, this gene was clustered into a separate category (Fig. 9). The remaining two categories included 9 genes with the highest expression in the expansion-formation period, and the expression of these genes tended to be downregulated overall; 18 genes had the highest expression in the early expansion period, and the changes in the expression patterns of these genes could be subdivided into four subclasses. Among these genes, *DoEXLA2* is particularly noteworthy. Its expression level

initially increased but then decreased, but the decrease began only in the late stage of swelling. Little difference in its expression level was observed between the early and middle stages of swelling. The expression levels of ten genes, namely, *DoEXPA1, DoEXPA3, DoEXPA4, DoEXPA5, DoEXPA8, DoEXPA10, DoEXPA12, DoEXPA16, DoEXPA18,* and *DoEXPB2,* decreased during the middle stage of tuber enlargement and reached their lowest levels in the late stage. The expression levels of three genes, *DoEXPA7*, *DoEXPA13*, and *DoEXPA20*, tended to first increase, then decrease, and then increase again, and the differences in their expression levels among the four stages were quite significant. The expression levels of four genes, *DoEXPA2, DoEXPA6, DoEXPA9*, and *DoEXPA17*, first tended to increase but then tended to decrease. The expression levels were similar and relatively high during the swelling formation stage and the early stage, and the expression levels were similar and relatively low during the middle and late stages of swelling. Moreover, among the 18 genes with high expression levels at the early stage of tuber swelling, 15 (83.3%) belong to the EXPA subfamily. The expression levels of the 3 genes in the *DoEXLB* subfamily tended to be relatively high during the tuber formation stage but then gradually decreased. Thus, the trends for the changes in the expression levels of genes within a subfamily are basically consistent.

GO functional enrichment analysis of *DoEXPs*

The functions of yam expansin genes were predicted by a GO functional enrichment analysis and consisted of three modules, namely, CC (cellular component), BP (biological process), and MF (molecular function). The *DoEXPs* included a total of 16 GO terms, including 1 cellular component and 15 biological processes, and were not enriched for molecular functions (Fig. S2). The information of the entries enriched in the cellular component term was plant-type

cell wall, which was consistent with the prediction that the yam EXP genes are localized in the cell wall in the early stage of this study; the biological processes were involved mainly in the growth and development of the plant root system and morphogenesis.

RNA-Seq data validation

To confirm the reliability of RNA-Seq data, a validation experiment on randomly selected *DoEXPs* were determined using qPCR (Fig.10). Expression patterns of the nine transcripts during tuber development were consistent with RNA Seq data despite the differential expression folds, which confirmed that our RNA-seq data and subsequent interpretations were reliable.

Discussion

Overview of the Cell Wall Composition and Expansins

The cell wall is a unique plant structure that plays key roles in plant growth, cell differentiation, intercellular communication, water movement, and defence (Cosgrove, 2005). It is primarily composed of three polysaccharides—cellulose, hemicellulose, and pectin—along with lignin and a small amount of cell wall proteins. Cellulose acts as the framework of the cell wall, consisting of layers of cellulose microfibrils (Zhong & Ye, 2015; Cosgrove, 2024a). Growing plant cells are formed by an extensible wall, which is a complex amalgam of cellulose microfilaments bonded noncovalently to a matrix of hemicellulose, pectin and structural proteins. The cell wall determines the shape of the plant cell. As the cell grows, the wall stretches to accommodate the expanding cell, a process critical for the overall development of the plant (Cosgrove, 1997). Expansins are a group of nonenzymatic cell wall proteins that can relieve cell wall stress and induce cell wall extension. These compounds are believed to disrupt the noncovalent bonds between cellulose microfibrils and play various biological roles. The main expansin families in plants include α-expansin (EXPA), which acts on cellulose–cellulose linkages; β-expansin (EXPB), which acts on xylan and is a major component of hemicellulose; and expansin-like A/B. EXPA mediates acidic growth, which contributes to the expansion of the cell wall by growth hormone and other growth factors (Cosgrove, 2024b). The auxin signal activates the plasma membrane H+-ATPase, thereby reducing the cell wall pH and further promoting cell wall growth (Du et al., 2020). EXPB is known as a Group 1 grass pollen allergen and is abundantly expressed in the pollen of gramineous plants. It may loosen the maternal cell

wall during the growth of the pollen tube towards the ovary (Sampedro and Cosgrove,2005). Expansin-like A/B has also been confirmed to be involved in seed and root development, as well as stress resistance processes, in various plants (Marowa et al., 2016; Muthusamy et al., 2020).

Identification and Functional Analysis of *DoEXPs*

In this study, through a whole-genome analysis, 30 expansin genes with two conserved domains were identified in yam. Similar to other plants, all the identified yam EXP genes (*DoEXPs*) were divided into four subfamilies: EXPA, EXPB, EXLA, and EXLB (Table 1). A total of 21, 2, 4, and 3 members were identified in each subfamily, respectively. Moreover, we analysed the chromosomal locations of the above genes and concluded that the EXP genes were located on 14 chromosomes and renamed the genes *DoEXPA1-21*, *DoEXPB1-2*, *DoEXLA1-4*, and *DoEXLB1-3*. An analysis of the conserved motifs and gene structures of *DoEXPs* revealed that members within the same subfamily presented a high degree of structural conservation. All *DoEXPs* contained two conserved motifs, 1 and 6, indicating that they played important roles in the evolution of the EXP gene family. An analysis of the structural motif results revealed that the numbers of introns and exons within each subfamily were consistent. The positions of the two conserved structural domains in the genes are consistent with the results of previous studies. By analysing the cis-acting elements in the 2000 bp promoter region upstream of DoEXPs, 20 of these genes were found to have light-responsive elements, suggesting that EXP genes are involved in the regulation of photomorphogenesis in plants. These findings are consistent with the results reported in wheat, where EXP genes respond to light signals to influence the developmental process of the plant (Yang et al., 2023). In addition, EXP genes are involved in multiple hormone response pathways, among which the number of response elements to abscisic acid (ABA) is as high as 37. When plants are challenged by pathogen attacks and abiotic stresses (e.g., drought, low temperature, and salt stress), abscisic acid serves as a key endogenous messenger to transmit stress signals to downstream pathways, prompting the plant to respond to the adverse external environment (Lee & Luan, 2012; Raghavendra et al., 2010). Based on these results, we hypothesized that EXP genes are involved in a variety of stress tolerance processes in plants, and when plants overexpress EXP genes, plant stress tolerance is increased, consistent with investigations in species such as cotton and wheat (Zhang et al., 2021; Chen et al., 2017). The EXP genes also have phytohormone response elements such as Auxin (IAA), gibberellin (GA), and methyl jasmonate (MeJA) response elements, among which, the most noteworthy are

elements that respond to growth hormone signalling. A total of 12 growth hormone response elements have been identified in the yam EXP genes, of which 8 are located in the EXPA subfamily of genes. As EXPA is regarded as a catalyst for 'acid growth', when the pH<5, growth hormone rapidly stimulates the plant cell wall through the acid growth mechanism to facilitate elongation (Durachko & Cosgrove, 2009). The results of the functional enrichment analysis of *DoEXPs* revealed that the DoEXPA subfamily is involved mainly in root growth and morphogenesis. By analysing the types and quantities of cis-acting elements contained in each gene, we found that four genes, *DoEXPA9, DoEXPA16, DoEXPA18*, and *DoEXPA20*, have ten or more acting elements. Moreover, *DoEXPA16* and *DoEXPA20* contain the most types of acting elements, with many important response elements, such as light-responsive, hormone-responsive, and abiotic stress-responsive elements, indicating that these two genes are involved in a wide range of biological processes. *DoEXPA18* has 13 response elements, with the greatest numbers of light-responsive, abscisic acid-responsive, and MeJA-responsive elements, indicating that this gene strongly responds to light, ABA, and MeJA signals. *DoEXPA9* has two drought-responsive elements and one low-temperature-responsive element, suggesting that this gene has a relatively strong ability to resist external abiotic stresses. The fact that the EXP genes are responsive to drought and low-temperature stress signals suggests that they can function under abiotic stress, a conclusion that has been confirmed in plants such as poplar, willow, and tobacco (Yin et al., 2023; Zhang et al., 2025; Marowa et al., 2020).

Evolution and Trends in the Expression of *DoEXPs*

The generation and evolution of gene families may be caused by tandem and segmental duplications (Kent et al., 2003). We studied the evolutionary relationships of the DoEXP gene family by analysing the duplication events of the DoEXP gene family. In this study, 2 pairs of tandem duplications and 16 pairs of segmental duplications were identified, indicating that the evolution of the DoEXP gene family occurred mainly through segmental duplication. This result is consistent with the evolutionary findings of the expansin gene family in multiple species reported by other researchers (Li et al., 2021; Han et al., 2019; Zhu et al., 2014). Through an analysis of the selection pressure of the duplicated gene pairs in *DoEXPs* and comparisons between *DoEXPs* and rice and Arabidopsis, we concluded that the DoEXP gene family underwent purifying selection during the evolutionary process. The trends of *DoEXP* expression in yam during four periods were analysed and divided into three major categories. Among them,

DoEXPA19 was the most unique, as its expression level was the highest in the late swelling stage. The genes in the second group presented the highest expression level during the tuber formation stage, and their expression levels gradually decreased, reaching the lowest expression level in the late swelling stage. The genes in the other groups presented the highest expression level in the early swelling stage and then gradually decreased. Given that the highest swelling rate of yam tubers occurs from the early to the middle swelling stage, the expression levels of yam expansin genes, which regulate tuber swelling, should peak during these stages. Based on the transcriptome gene expression data, we speculated that the 18 genes with the highest expression levels in the early swelling stage play important roles in regulating the swelling of yam tubers. Moreover, DoEXPA9, DoEXPA16, DoEXPA18, and DoEXPA20 contained relatively large numbers and varieties of cis-acting elements, and their expression levels were highest in the early stage of yam tuber swelling. This characteristic was also verified by qRT–PCR, which suggested that these four genes could be used as the main entry points to verify the impact of DoEXPs on yam tuber swelling, as well as whether they affect the hormone response and tolerance of transgenic plants to abiotic stress.

Conclusions

Overall, this study identified and analysed the genes of the expansin family in yam at the whole-genome level for the first time, identifying 30 expansin genes and analysing them from the perspectives of their physicochemical properties, evolutionary relationships, and biological functions. The results revealed that expansin genes play important roles in the response to abiotic stress and in the growth and development of plant roots. These findings advance understanding of expansin functions in yam, providing a foundation for future research on tuber development and stress adaptation, with potential applications in breeding high-yield, stress-resistant varieties. Further experimental validation of gene functions is needed to confirm these insights.

References:

Abbasi, A., Malekpour, M., and Sobhanverdi, S. 2021. The Arabidopsis expansin gene (AtEXPA18) is

capable to ameliorate drought stress tolerance in transgenic tobacco plants. MOLECULAR BIOLOGY REPORTS 48:5913-5922. 10.1007/s11033-021-06589-2

Cannon, S.B., Mitra, A., Baumgarten, A., Young, N.D., and May, G. 2004. The roles of segmental and tandem gene duplication in the evolution of large gene families in Arabidopsis thaliana. BMC PLANT BIOLOGY 4:10. 10.1186/1471-2229-4-10

Chang, Y., Sun, H., Liu, S., He, Y., Zhao, S., Wang, J., Wang, T., Zhang, J., Gao, J., Yang, Q., Li, M., and Zhao, X. 2023. Identification of BBX gene family and its function in the regulation of microtuber formation in yam. BMC GENOMICS 24:354. 10.1186/s12864-023-09406-1

Che, J., Yamaji, N., Shen, R.F., and Ma, J.F. 2016. An Al-inducible expansin gene, OsEXPA10 is involved in root cell elongation of rice. PLANT JOURNAL 88:132-142. 10.1111/tpj.13237

Chen, Y., Han, Y., Kong, X., Kang, H., Ren, Y., and Wang, W. 2017. Ectopic expression of wheat expansin gene TaEXPA2 improved the salt tolerance of transgenic tobacco by regulating Na(+) /K(+) and antioxidant competence. PHYSIOLOGIA PLANTARUM 159:161-177. 10.1111/ppl.12492

Chen, Y., Han, Y., Zhang, M., Zhou, S., Kong, X., and Wang, W. 2016. Overexpression of the Wheat Expansin Gene TaEXPA2 Improved Seed Production and Drought Tolerance in Transgenic Tobacco Plants. PLoS One 11:e153494. 10.1371/journal.pone.0153494

Choi, D., Lee, Y., Cho, H., and Kende, H. 2003. Regulation of expansin gene expression affects growth and development in transgenic rice plants. PLANT CELL 15:1386-1398. 10.1105/tpc.011965

Cosgrove, D.J. 1997. Assembly and enlargement of the primary cell wall in plants. Annual Review of Cell and Developmental Biology 13:171-201. 10.1146/annurev.cellbio.13.1.171

Cosgrove, D.J. 2000. Loosening of plant cell walls by expansins. NATURE 407:321-326. 10.1038/35030000

Cosgrove, D.J. 2005. Growth of the plant cell wall. NATURE REVIEWS MOLECULAR CELL BIOLOGY 6:850-861. 10.1038/nrm1746

Cosgrove, D.J. 2024a. Structure and growth of plant cell walls. NATURE REVIEWS MOLECULAR CELL BIOLOGY 25:340-358. 10.1038/s41580-023-00691-y

Cosgrove, D.J. 2024b. Plant Cell Wall Loosening by Expansins. Annual Review of Cell and Developmental Biology 40:329-352. 10.1146/annurev-cellbio-111822-115334

Du, M., Spalding, E.P., and Gray, W.M. 2020. Rapid Auxin-Mediated Cell Expansion. Annual Review of Plant Biology 71:379-402. 10.1146/annurev-arplant-073019-025907

Durachko, D.M., and Cosgrove, D.J. 2009. Measuring plant cell wall extension (creep) induced by acidic pH and by alpha-expansin. Jove-Journal of Visualized Experiments:1263. 10.3791/1263

Gao, W., Li, D., Fan, X., Sun, Y., Han, B., Wang, X., and Xu, G. 2020. Genome-wide identification, characterization, and expression analysis of the expansin gene family in watermelon (Citrullus lanatus). 3 Biotech 10:302. 10.1007/s13205-020-02293-3

Han, Z., Liu, Y., Deng, X., Liu, D., Liu, Y., Hu, Y., and Yan, Y. 2019. Genome-wide identification and expression analysis of expansin gene family in common wheat (Triticum aestivum L.). BMC GENOMICS 20:101. 10.1186/s12864-019-5455-1

He, X., Zeng, J., Cao, F., Ahmed, I.M., Zhang, G., Vincze, E., and Wu, F. 2015. HvEXPB7, a novel beta-expansin gene revealed by the root hair transcriptome of Tibetan wild barley, improves root hair growth under drought stress. JOURNAL OF EXPERIMENTAL BOTANY 66:7405-7419. 10.1093/jxb/erv436

606 Jin, K., Zhuo, R., Xu, D., Wang, Y., Fan, H., Huang, B., and Qiao, G. 2020. Genome-Wide Identification
607 of the Expansin Gene Family and Its Potential Association with Drought Stress in Moso Bamboo.
608 INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 21. 10.3390/ijms21249491
609 Kende, H., Bradford, K., Brummell, D., Cho, H., Cosgrove, D., Fleming, A., Gehring, C., Lee, Y.,
610 McQueen-Mason, S., Rose, J., and Voesenek, L.A.C.J. 2004. Nomenclature for members of the expansin
611 superfamily of genes and proteins. PLANT MOLECULAR BIOLOGY 55:311-314. 10.1007/s11103-004-
612 0158-6
613 Kent, W.J., Baertsch, R., Hinrichs, A., Miller, W., and Haussler, D. 2003. Evolution's cauldron:
614 duplication, deletion, and rearrangement in the mouse and human genomes. PROCEEDINGS OF THE
615 NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 100:11484-11489.
616 10.1073/pnas.1932072100
617 Lee, S.C., and Luan, S. 2012. ABA signal transduction at the crossroad of biotic and abiotic stress
618 responses. PLANT CELL AND ENVIRONMENT 35:53-60. 10.1111/j.1365-3040.2011.02426.x
619 Li, J., Liu, L., Wang, L., Rao, I.M., Wang, Z., and Chen, Z. 2024. AcEXPA1, an alpha-expansin gene,
620 participates in the aluminum tolerance of carpetgrass (Axonopus compressus) through root growth regulation.
621 PLANT CELL REPORTS 43:159. 10.1007/s00299-024-03243-6
622 Li, K., Ma, B., Shen, J., Zhao, S., Ma, X., Wang, Z., Fan, Y., Tang, Q., and Wei, D. 2021. The evolution
623 of the expansin gene family in Brassica species. PLANT PHYSIOLOGY AND BIOCHEMISTRY 167:630-
624 638. 10.1016/j.plaphy.2021.08.033
625 Lin, Z., Ni, Z., Zhang, Y., Yao, Y., Wu, H., and Sun, Q. 2005. Isolation and characterization of 18 genes
626 encoding alpha- and beta-expansins in wheat (Triticum aestivum L.). MOLECULAR GENETICS AND
627 GENOMICS 274:548-556. 10.1007/s00438-005-0029-0
628 Livak, K.J., and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time
629 quantitative PCR and the 2(-Delta Delta C(T)) Method. METHODS 25:402-408. 10.1006/meth.2001.1262
630 Marga, F., Grandbois, M., Cosgrove, D.J., and Baskin, T.I. 2005. Cell wall extension results in the
631 coordinate separation of parallel microfibrils: evidence from scanning electron microscopy and atomic force
632 microscopy. PLANT JOURNAL 43:181-190. 10.1111/j.1365-313X.2005.02447.x
633 Marowa, P., Ding, A., Xu, Z., and Kong, Y. 2020. Overexpression of NtEXPA11 modulates plant growth
634 and development and enhances stress tolerance in tobacco. PLANT PHYSIOLOGY AND BIOCHEMISTRY
635 151:477-485. 10.1016/j.plaphy.2020.03.033
636 Marowa, P., Ding, A., and Kong, Y. 2016. Expansins: roles in plant growth and potential applications in
637 crop improvement. PLANT CELL REPORTS 35:949-965. 10.1007/s00299-016-1948-4
638 Muthusamy, M., Kim, J.Y., Yoon, E.K., Kim, J.A., and Lee, S.I. 2020. BrEXLB1, a Brassica rapa
639 Expansin-Like B1 Gene is Associated with Root Development, Drought Stress Response, and Seed
640 Germination. Genes 11. 10.3390/genes11040404
641 Obidiegwu, J.E., Lyons, J.B., and Chilaka, C.A. 2020. The Dioscorea Genus (Yam)-An Appraisal of
642 Nutritional and Therapeutic Potentials. Foods 9. 10.3390/foods9091304
643 Raghavendra, A.S., Gonugunta, V.K., Christmann, A., and Grill, E. 2010. ABA perception and signalling.
644 TRENDS IN PLANT SCIENCE 15:395-401. 10.1016/j.tplants.2010.04.006
645 Sampedro, J., and Cosgrove, D.J. 2005. The expansin superfamily. GENOME BIOLOGY 6:242.
646 10.1186/gb-2005-6-12-242

Shcherban, T.Y., Shi, J., Durachko, D.M., Guiltinan, M.J., McQueen-Mason, S.J., Shieh, M., and Cosgrove, D.J. 1995. Molecular cloning and sequence analysis of expansins--a highly conserved, multigene family of proteins that mediate cell wall extension in plants. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 92:9245-9249. 10.1073/pnas.92.20.9245

Shi, Y., Xu, X., Li, H., Xu, Q., and Xu, J. 2014. [Bioinformatics analysis of the expansin gene family in rice]. Yi Chuan 36:809-820. 10.3724/SP.J.1005.2014.0809

Sun, W., Yao, M., Wang, Z., Chen, Y., Zhan, J., Yan, J., Jiang, S., Jian, S., Chen, H., Bu, T., Tang, Z., Li, Q., Zhao, H., and Wu, Q. 2022. Involvement of Auxin-Mediated CqEXPA50 Contributes to Salt Tolerance in Quinoa (Chenopodium quinoa) by Interaction with Auxin Pathway Genes. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 23. 10.3390/ijms23158480

Wang, L., Zhang, T., Li, C., Zhou, C., Liu, B., Wu, Y., He, F., Xu, Y., Li, F., and Feng, X. 2024. Overexpression of Wild Soybean Expansin Gene GsEXLB14 Enhanced the Tolerance of Transgenic Soybean Hairy Roots to Salt and Drought Stresses. Plants-Basel 13. 10.3390/plants13121656

Wang, X., Ma, J., He, F., Wang, L., Zhang, T., Liu, D., Xu, Y., Li, F., and Feng, X. 2024. A Study on the Functional Identification of Overexpressing Winter Wheat Expansin Gene TaEXPA7-B in Rice under Salt Stress. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 25. 10.3390/ijms25147707

Wu, Y., Meeley, R.B., and Cosgrove, D.J. 2001. Analysis and expression of the alpha-expansin and beta-expansin gene families in maize. PLANT PHYSIOLOGY 126:222-232. 10.1104/pp.126.1.222

Wu, Z., Li, M., Zhong, Y., Li, L., Cheng, D., Gu, H., Guo, X., Qi, X., and Chen, J. 2022. Overexpression of AcEXPA23 Promotes Lateral Root Development in Kiwifruit. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 23. 10.3390/ijms23148026

Yang, H., Li, Y., Qiao, Y., Sun, H., Liu, W., Qiao, W., Li, W., Liu, M., and Dong, B. 2023. Low light stress promotes new tiller regeneration by changing source-sink relationship and activating expression of expansin genes in wheat. PLANT CELL AND ENVIRONMENT 46:1562-1581. 10.1111/pce.14548

Yin, Z., Zhou, F., Chen, Y., Wu, H., and Yin, T. 2023. Genome-Wide Analysis of the Expansin Gene Family in Populus and Characterization of Expression Changes in Response to Phytohormone (Abscisic Acid) and Abiotic (Low-Temperature) Stresses. INTERNATIONAL JOURNAL OF MOLECULAR SCIENCES 24. 10.3390/ijms24097759

Zhang, B., Chang, L., Sun, W., Ullah, A., and Yang, X. 2021. Overexpression of an expansin-like gene, GhEXLB2 enhanced drought tolerance in cotton. PLANT PHYSIOLOGY AND BIOCHEMISTRY 162:468-475. 10.1016/j.plaphy.2021.03.018

Zhang, J., Wang, L., Zhao, H., Gong, L., and Xu, J. 2025. The SmWRKY12-SmRAP2-7-SmEXPA13 module in Salix matsudana koidz enhances plant tolerance to drought stress. INTERNATIONAL JOURNAL OF BIOLOGICAL MACROMOLECULES 284:138077. 10.1016/j.ijbiomac.2024.138077

Zhong, R., and Ye, Z. 2015. Secondary cell walls: biosynthesis, patterned deposition and transcriptional regulation. PLANT AND CELL PHYSIOLOGY 56:195-214. 10.1093/pcp/pcu140

Zhou, Y., Li, Y., Gong, M., Qin, F., Xiao, D., Zhan, J., Wang, A., and He, L. 2021. Regulatory mechanism of GA(3) on tuber growth by DELLA-dependent pathway in yam (Dioscorea opposita). PLANT MOLECULAR BIOLOGY 106:433-448. 10.1007/s11103-021-01163-7

Zhu, Y., Wu, N., Song, W., Yin, G., Qin, Y., Yan, Y., and Hu, Y. 2014. Soybean (Glycine max) expansin

688 gene superfamily origins: segmental and tandem duplication events followed by divergent selection among
689 subfamilies. BMC PLANT BIOLOGY 14:93. 10.1186/1471-2229-14-93

Figure 1

Phylogenetic evolutionary tree analysis

The phylogenetic tree is constructed using expansin proteins from several species, including *O . sativa , A. thaliana , D . opposite ,* and *T . aestivum ,* in the figure, Os represents rice, At represents Arabidopsis , Do represents yam, Ta represents wheat. EXPA , EXPB, EXLA, and EXLB are the names of the four subfamilies of the EXP family.

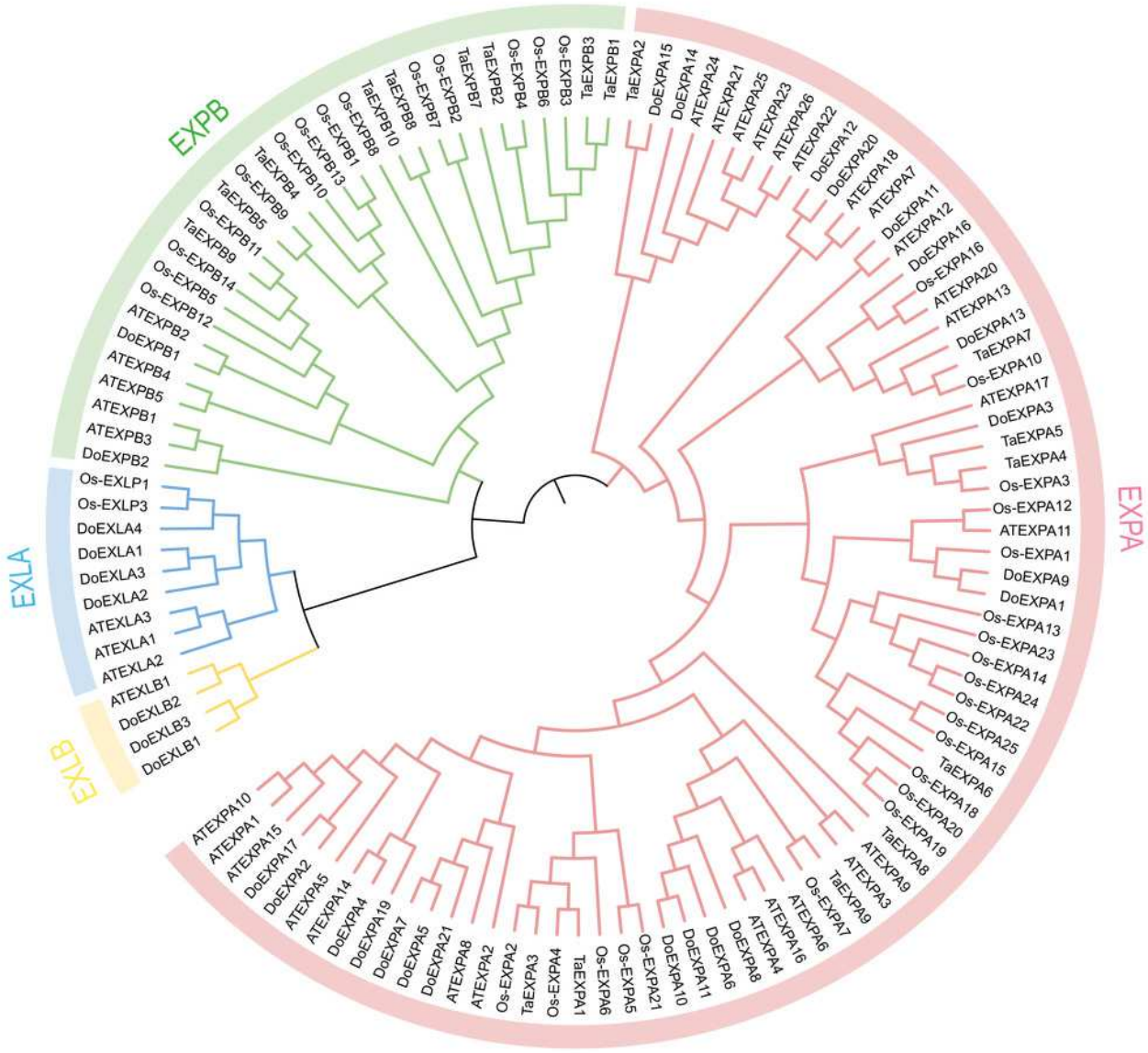

Figure 2

Phylogenetic Relationships, Motif Composition, stable structural domains, and Gene Structures of *DoEXPs*

(A) The phylogenetic relationships of *DoEXPs* , categorized into four groups: α-expansin (EXPA), β-expansin (EXPB), expansin-like A (EXLA), and expansin-like B (EXLB), represented by pink, green, yellow, and blue, respectively; (C) Interpro database to unlock the conserved structural domains and positional information of *DoEXPs* ; (D) Structural analysis of mountain *DoEXPs* , with intron and exon structures shown.

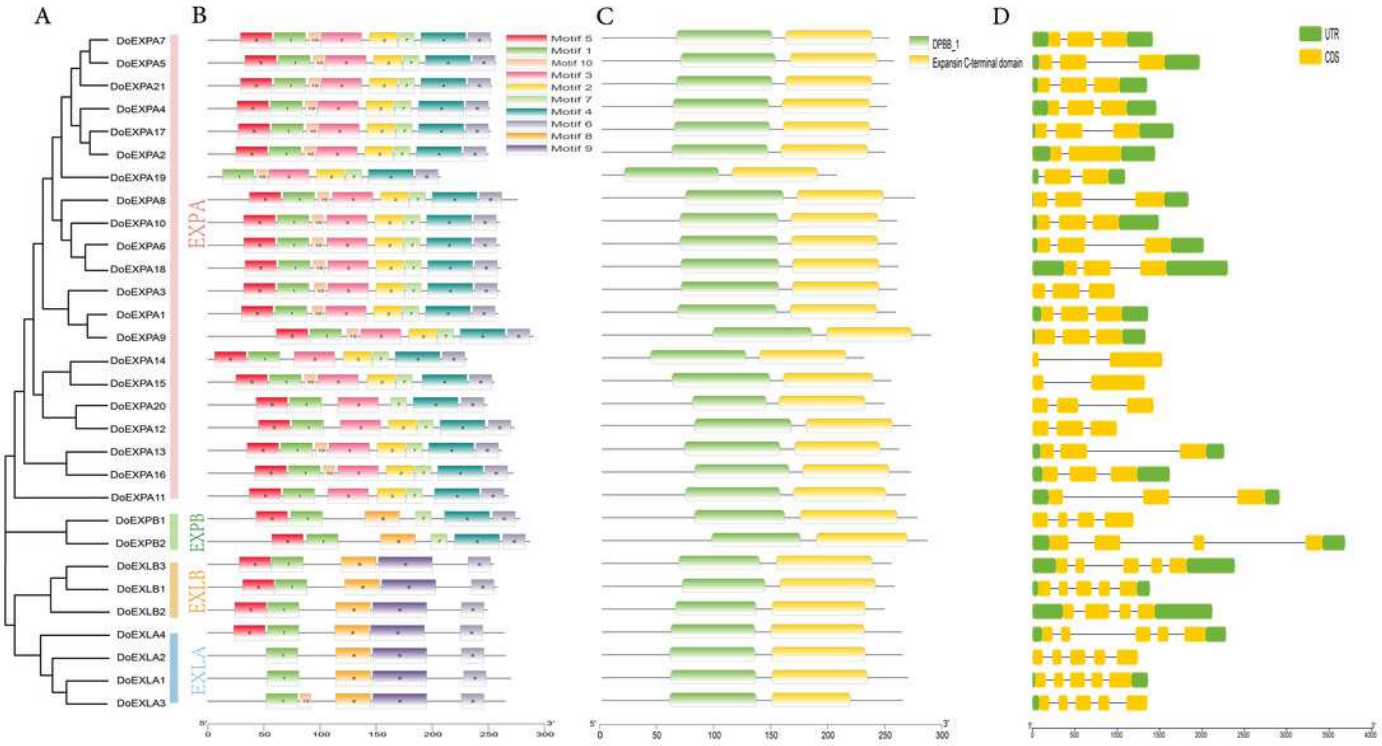

Figure 3

Chromosome mapping diagram of *DoEXPs* .

Gene density is displayed on chromosomes in the form of a heat map, with high gene density indicated in red and low gene density indicated in blue. Two pairs of tandem duplication gene pairs are represented by blue lines.

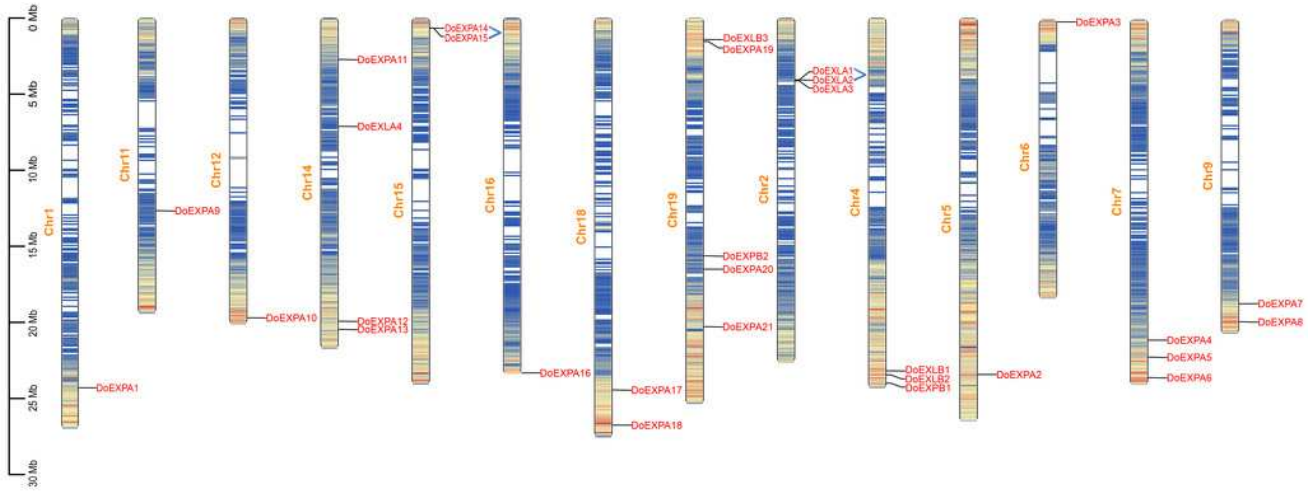

Figure 4

Cis-acting element analysis of DoEXPs

The Plantcare online website was used to predict the cis-acting elements in the promoters 2000bp upstream of the Do EXP genes. Different colored circles were used to represent different cis-acting elements .

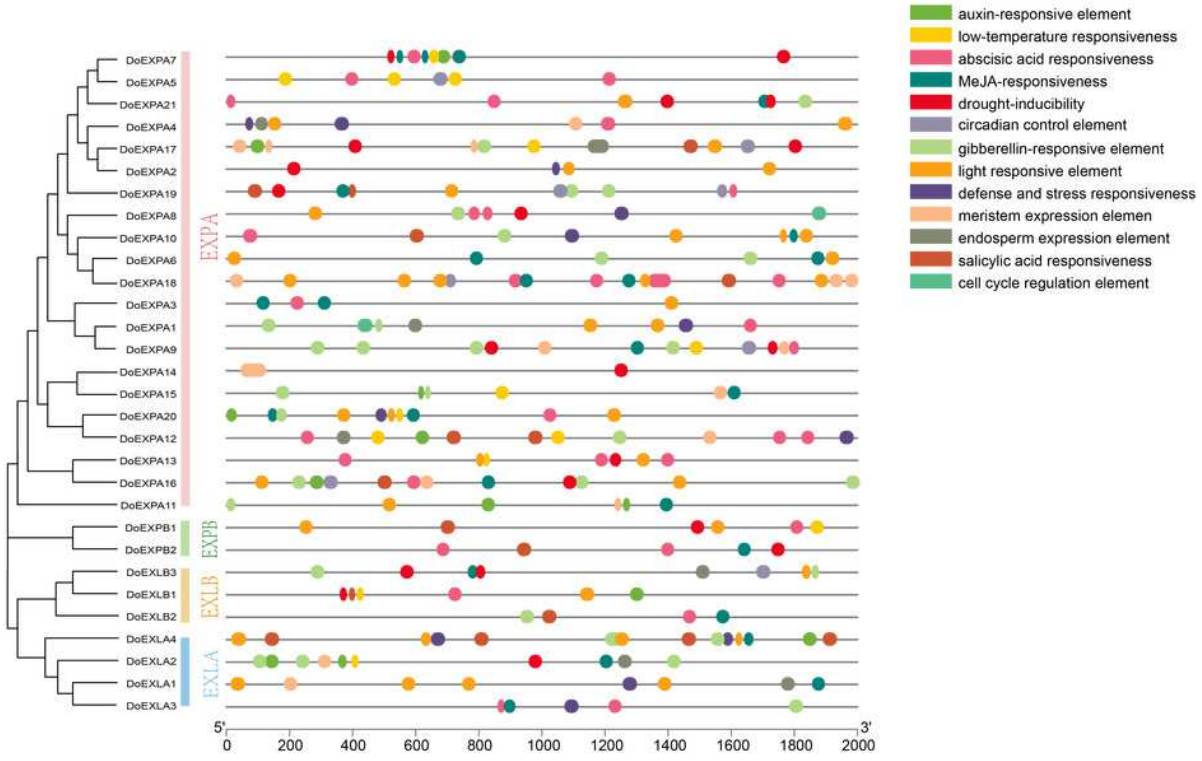

Figure 5

Heat map of the number of cis-acting elements in the promoter regions of DoEXPs.

Seven important cis-acting elements of DoEXPs are represented by different colored boxes with numbers

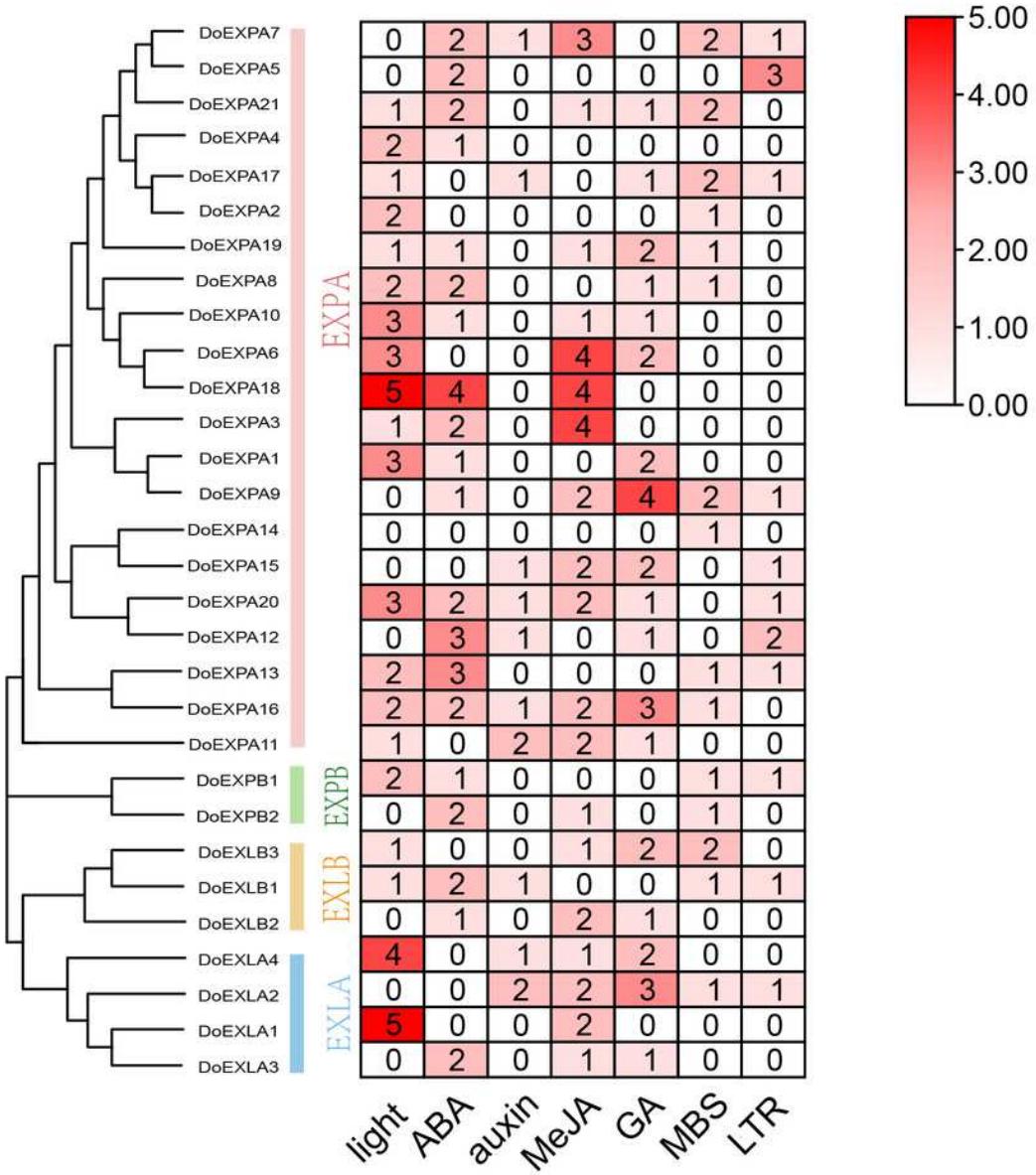

Figure 6

Intraspecific collinearity analysis of *DoEXPs* .

From the outside to the inside in the figure, there are gene names, chromosome names, a line graph of gene density, a heat map of gene density, and the collinearity situation among genes in sequence. The gray lines in the figure refer to the gene collinearity information within the whole genome of yam.

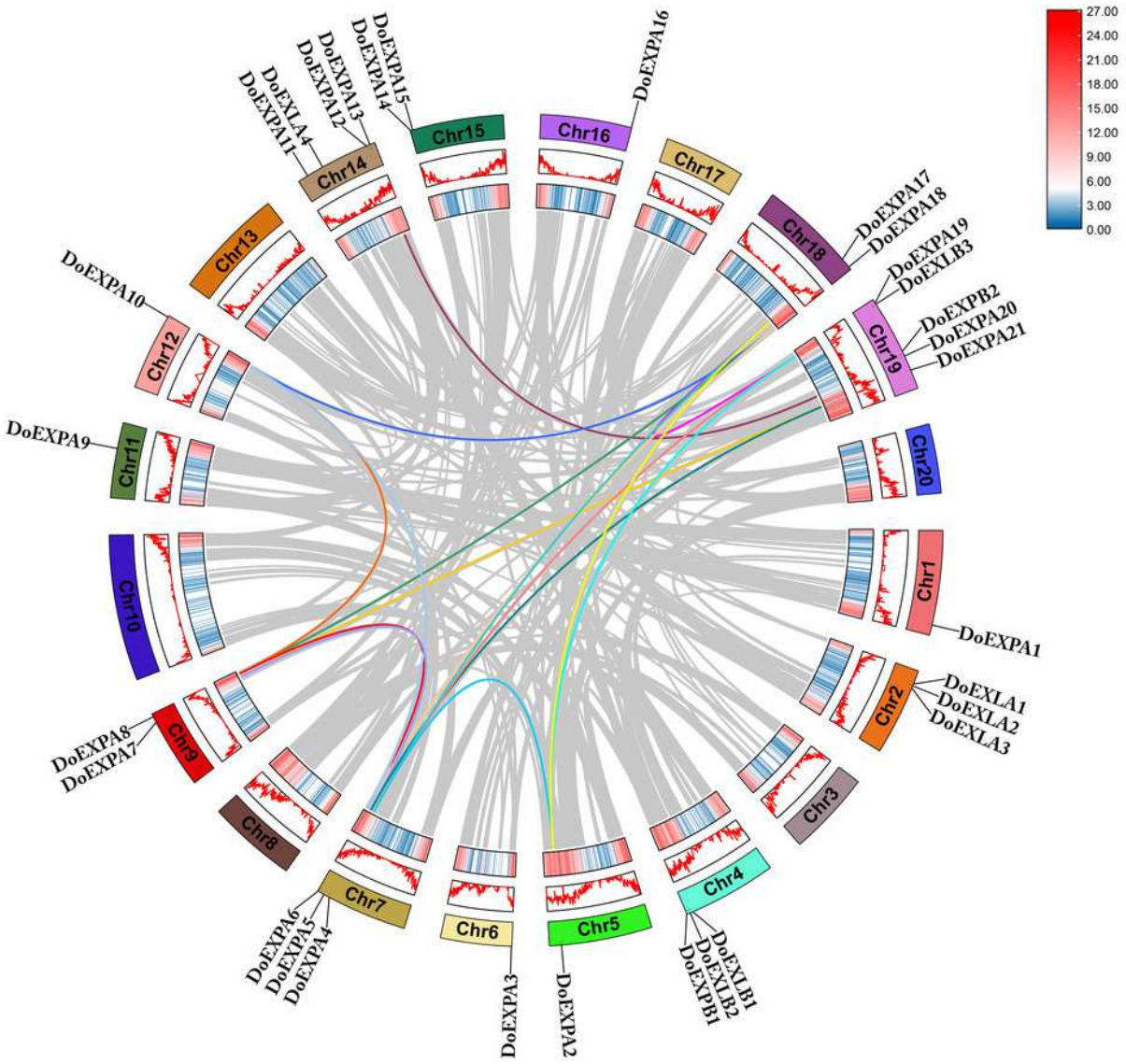

Figure 7

Synteny analyses of *DoEXP s* to Arabidopsis and rice .

The grey lines in the background indicate all the covariance information in the genomes of yam with Arabidopsis and rice, a nd different color blocks represent the chromosomes of different species.

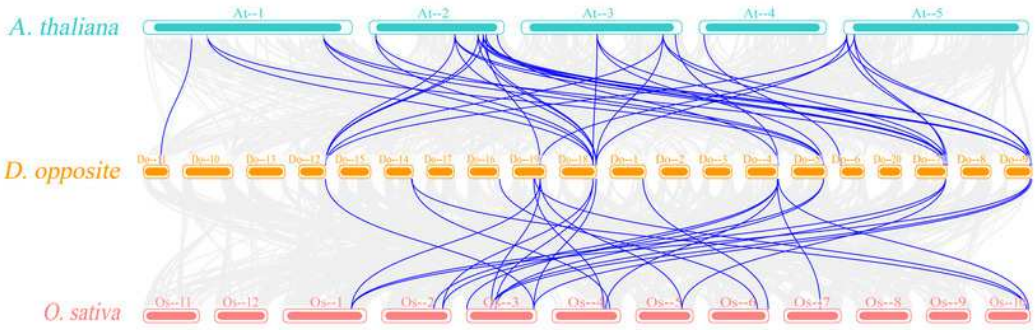

Figure 8

The Ka/Ks values of all collinear gene pairs within and between species

Distribution of Ka/Ks of all homologous gene pairs within yam (Do-Do), between yam and rice (Os-Do), and between yam and Arabidopsis thaliana (AT-Do). They are represented by green triangles, blue squares and pink circles respectively, and the red dotted line indicates the slope where Ka/Ks = 1.

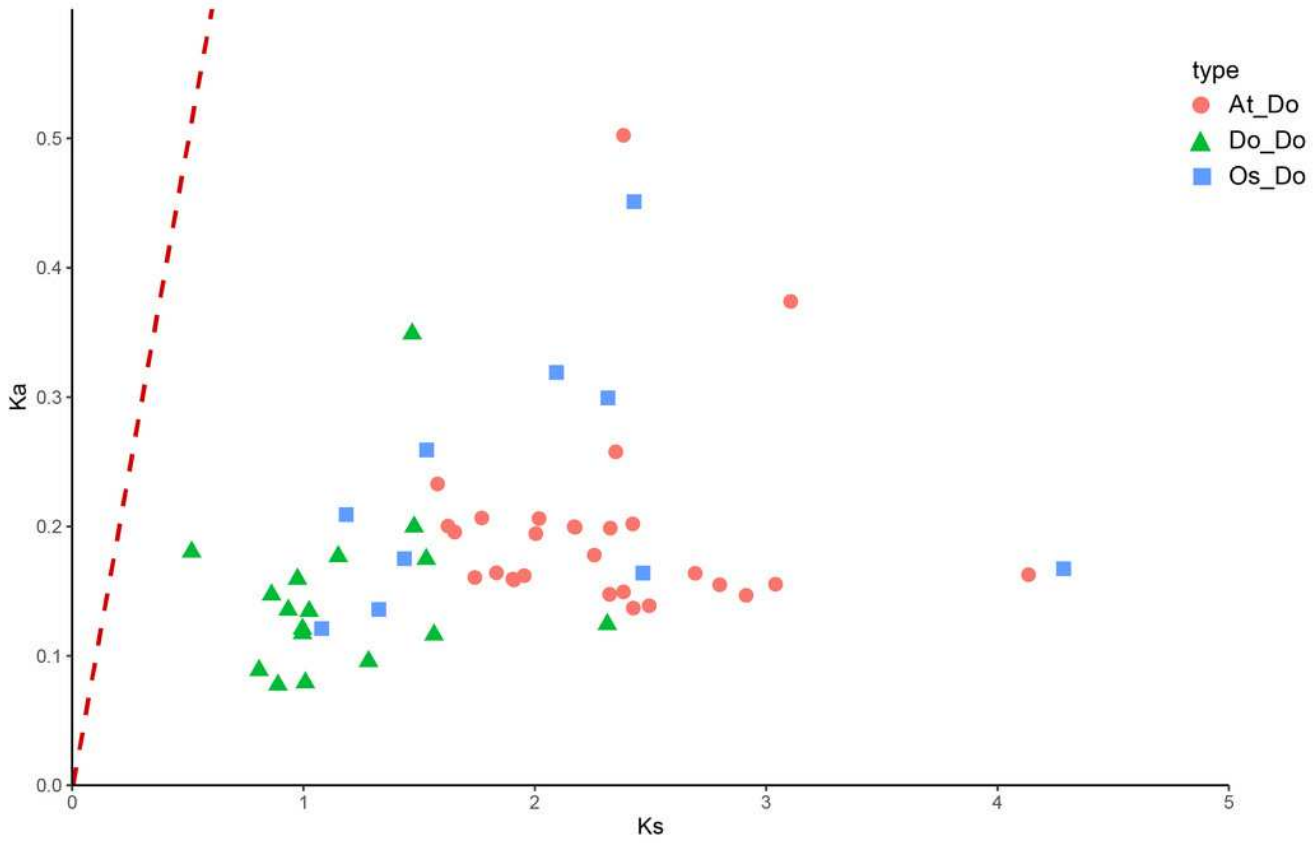

Figure 9

Heat map analysis of the expression of *DoEXPs*

The abscissa NH_F/E/M/L represents the four periods of variety NH1 (abbreviated as NH), namely F: tuber formation stage, E:early tuber expansion, M: mid-tuber expansion, and L: late tuber expansion .

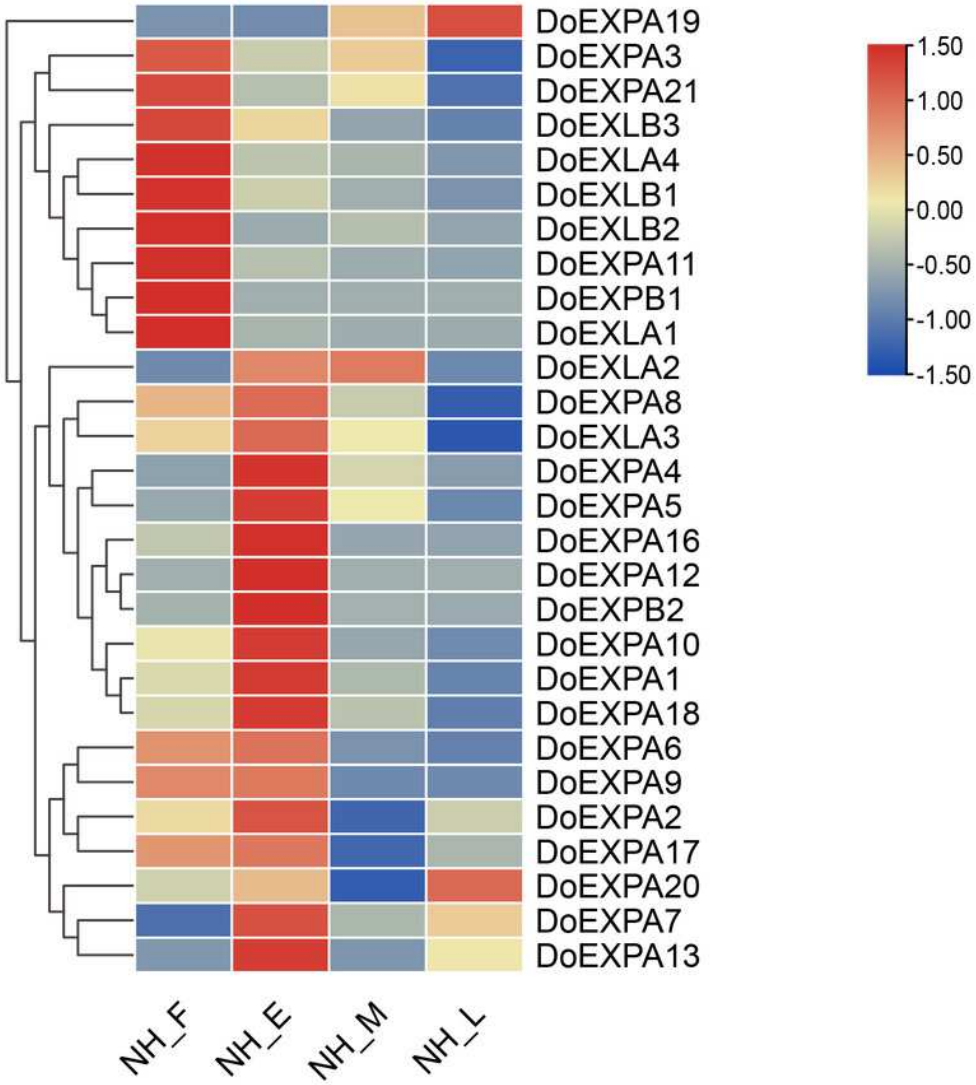

Figure 10

qPCR expression analysis of 9 selected *DoEXPs* in different period s.

The x-axis shows the different periods of tuber development, the y-axis shows the relative expression of the genes, and the data are expressed as SD ± mean.

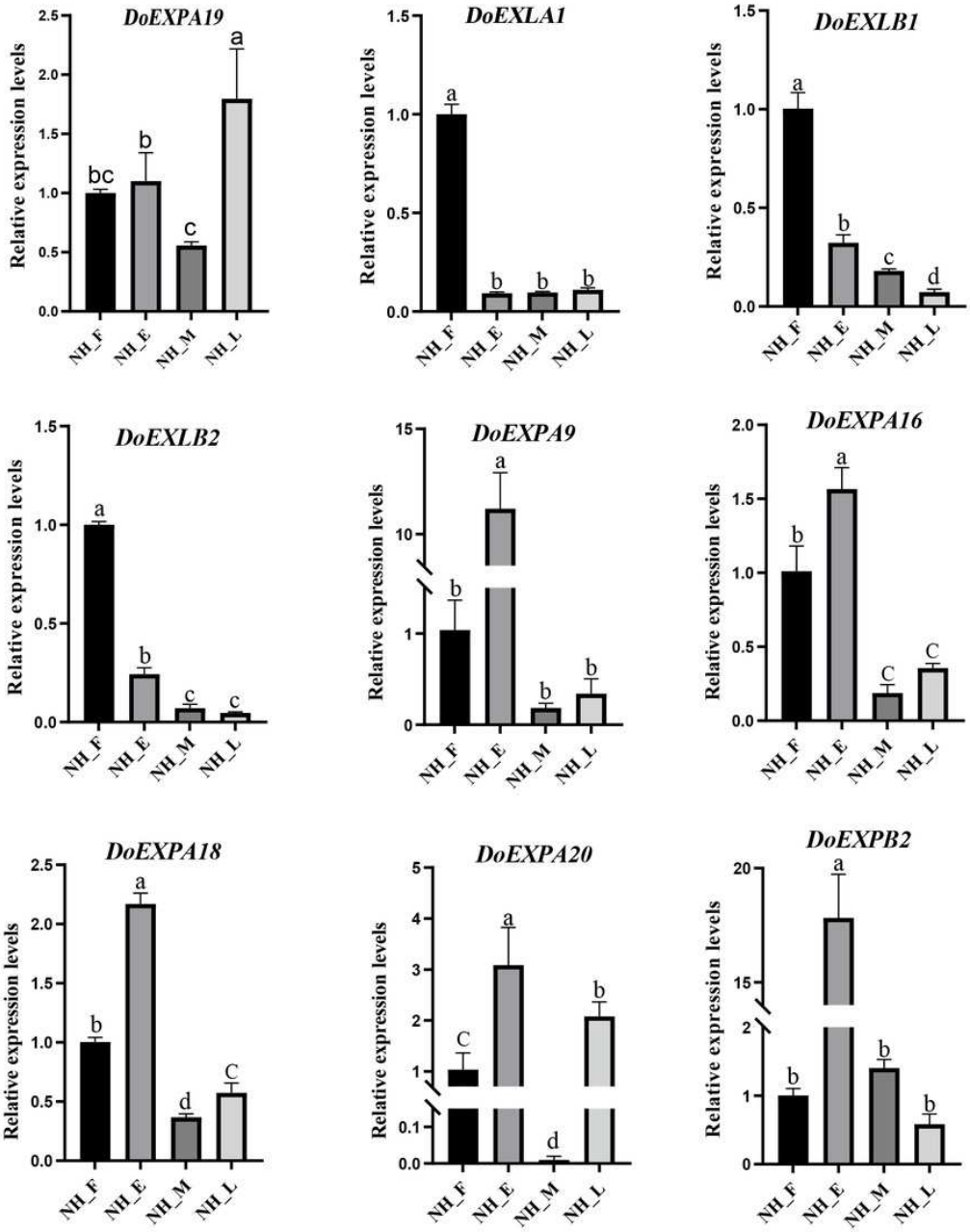
DoEXPA19
DoEXLA1
DoEXLB1
DoEXLB2
DoEXPA9
DoEXPA16
DoEXPA18
DoEXPA20
DoEXPB2
Relative expression levels
NH_F
NH_E
NH_M
NH_L

Table 1(on next page)

Details of the identified expansin genes in yam

gene ID	gene name	start	end	strand	length (aa)	PI	MW (Da)	hydrophilicity	Subcellular location
Dioal.01G060200.1.p	DoEXPA1	24298694	24300058	−	258	8.54	28087.72	-0.06	cell wall
Dioal.05G195500.1.p	DoEXPA2	23425696	23427143	−	249	8.12	26906.08	-0.077	cell wall
Dioal.06G001700.1.p	DoEXPA3	128847	129817	−	259	9.37	28480.21	-0.244	cell wall
Dioal.07G093100.1.p	DoEXPA4	21043371	21045031	−	250	7.52	26492.72	-0.026	cell wall
Dioal.07G106800.1.p	DoEXPA5	22165321	22167293	−	256	6.86	27441.79	0.034	cell wall
Dioal.07G127900.1.p	DoEXPA6	23506033	23508052	−	259	9.45	27882.7	-0.046	cell wall
Dioal.09G059800.1.p	DoEXPA7	18651741	18653159	−	252	8.12	26823.91	-0.087	cell wall
Dioal.09G077500.1.p	DoEXPA8	19835585	19841937	−	275	8.06	30361.68	0.056	cell wall
Dioal.11G043900.1.p	DoEXPA9	12662460	12663794	+	289	8.78	31599.85	-0.017	cell wall
Dioal.12G089000.1.p	DoEXPA10	19701720	19703208	−	259	9.32	27804.85	0.055	cell wall
Dioal.14G031400.1.p	DoEXPA11	2723362	2726278	−	267	9.66	28910.06	0.024	cell wall
Dioal.14G121100.1.p	DoEXPA12	19925968	19926962	+	272	9.17	29188.13	-0.107	cell wall
Dioal.14G128000.1.p	DoEXPA13	20458705	20460966	−	261	8.59	28028.25	0.129	cell wall
Dioal.15G009800.1.p	DoEXPA14	662455	663986	+	230	6.8	24822.82	-0.243	cell wall
Dioal.15G009900.1.p	DoEXPA15	665820	667145	+	254	9.44	27870.21	-0.104	cell wall
Dioal.16G093500.1.p	DoEXPA16	23322610	23324229	+	271	8.05	29706.79	0.093	cell wall
Dioal.18G071400.1.p	DoEXPA17	24445322	24446987	−	251	9.11	26627	0.024	cell wall
Dioal.18G109200.1.p	DoEXPA18	26752627	26754932	−	260	9.34	27853.68	-0.007	cell wall
Dioal.19G023600.1.p	DoEXPA19	1540636	1541727	+	206	8.61	22016.99	0.011	cell wall
Dioal.19G082500.1.p	DoEXPA20	16501589	16503014	+	248	9.3	26617.22	0.002	cell wall
Dioal.19G119800.1.p	DoEXPA21	20283831	20285183	−	252	8.4	27128.46	-0.098	cell wall
Dioal.04G180000.1.p	DoEXPB1	23997238	23998430	+	277	6.41	29230.91	-0.049	cell wall
Dioal.19G078100.1.p	DoEXPB2	15631278	15634965	−	286	8.05	31344.75	-0.1	cell wall
Dioal.02G029200.1.p	DoEXLA1	4063537	4064899	−	269	7.99	29150.32	-0.025	cell wall
Dioal.02G029300.1.p	DoEXLA2	4067080	4068322	−	264	9.07	28691.12	-0.046	cell wall
Dioal.02G029800.1.p	DoEXLA3	4128602	4130133	+	264	9	28604.88	0.046	cell wall
Dioal.14G049100.1.p	DoEXLA4	7116372	7118655	−	263	8.47	28761.91	-0.011	cell wall

Dioal.04G166400.1.p	DoEXLB1	23181553	23182937	−	257	4.82	27356.81	0.004	cell wall
Dioal.04G171000.1.p	DoEXLB2	23445733	23447854	+	248	7.43	27458.34	-0.17	cell wall
Dioal.19G021600.1.p	DoEXLB3	1411885	1414271	+	254	4.43	27227.49	-0.07	cell wall

1

Table 2(on next page)

Details of the identified expansin genes in yam

Seq_1	Seq_2	Ka	Ks	ka/ks	Date(Mya)
DoEXPA14	DoEXPA15	0.348976301	1.465376447	0.238147885	11.2721
DoEXLA1	DoEXLA2	0.180427892	0.512222507	0.352245148	3.9402
DoEXPA17	DoEXPA2	0.117321185	0.991671406	0.118306512	7.6282
DoEXPA19	DoEXPA2	0.147238695	0.857937152	0.171619441	6.5995
DoEXPA2	DoEXPA4	0.174745763	1.526510634	0.11447399	11.7424
DoEXPA17	DoEXPA4	0.199822805	1.474019019	0.135563247	11.3386
DoEXPA21	DoEXPA5	0.124372587	2.309011017	0.053864008	17.7616
DoEXPA5	DoEXPA7	0.095810531	1.277029241	0.075026106	9.8233
DoEXPA10	DoEXPA6	0.088856737	0.803862941	0.110537172	6.1836
DoEXPA18	DoEXPA6	0.077569824	0.88546908	0.087603086	6.8113
DoEXPA6	DoEXPA8	0.134544473	1.020308806	0.131866424	7.8485
DoEXPA21	DoEXPA7	0.116288398	1.559924484	0.074547453	11.9994
DoEXPA10	DoEXPA8	0.135464572	0.929768165	0.14569715	7.1521
DoEXPA18	DoEXPA8	0.120814847	0.991560743	0.121843112	7.6274
DoEXPA10	DoEXPA18	0.079365162	1.004022715	0.079047178	7.7233
DoEXPA12	DoEXPA20	0.176894509	1.146131211	0.154340539	8.8164
DoEXPA17	DoEXPA19	0.159389593	0.969968684	0.164324473	7.4613

1