

Genome-Wide Characterization of Expansin Gene Family in Cannabis sativa

Hülya SİPAHİ¹ 🗖

¹ Eskisehir Osmangazi University, Faculty of Agriculture, Department of Agricultural Biotechnology, Eskisehir *Corresponding author: hulya.sipahi@ogu.edu.tr

Abstract

Expansins loose plant cell walls and play diverse roles in plant growth and development, germination, fruit ripening and softening, fiber development, and biotic/abiotic stress response, especially adaptation to the osmotic and oxidative stresses caused by drought stress. In this study, genome-wide analysis of the expansin gene family in Cannabis sativa was performed for the first time. A total of 29 expansin genes were found in the C. sativa genome. These genes were classified into four subfamilies, including 18 CsEXPAs, 5 CsEXPBs, 1 CsEXLAs, and 5 CsEXLBs family members. Phylogenetic analysis showed that the cannabis, Arabidopsis and rice expansins were divided into ten subgroups. 29 cannabis expansin genes were unevenly distributed among nine cannabis chromosomes. Most expansin genes have 3 exons while the number of introns and exons among expansin genes ranged from 1 to 4 and from 2 to 5, respectively. The promoter regions of 29 cannabis expansin genes contained diverse cis-elements that are involved in the development, environmental stress, hormones, and light responsiveness. This initial study is a useful resource for further research on the potential roles of expansin in fiber development, plant growth and development, and environmental stress response.

Research Article

Article History	
Received	:30.11.2023
Accepted	:30.12.2023

Keywords Expansin Cannabis sativa bioinformatics

1. Introduction

The plant cell wall provides structural integrity and physical support to the plant, and the effects of cell wall components have important implications in many areas, from the cell's response to environmental stimuli to the production of renewable fuels, from fiber quality for textile and paper production to food quality and processing. Plant cell wall contain cellulose microfibrils, hemicellulose, pectin, lignin, and protein. During plant growth or adaptation changing environmental to conditions, the cell wall can be loose, remodeled, and elongated. Expansin proteins located in the cell wall can rearrange the shape of cellulose microfibrils in the extracellular matrix by loosening the H-bonds in a pHdependent manner (Sampedro and Cosgrove, 2005).

The plant expansin multigene family (EXP) consists of α -expansin (EXPA), β -expansin (EXPB), expansin-like A (EXLA) and expansin-like В (EXLB) subfamilies according phylogenetic relationship to (Sampedro and Cosgrove, 2005). Plant expansing are small proteins of approximately 250-275 amino acids in size, and they contain Pollen_allerg_1 DPBB 1 and conserved domains (Sampedro and Cosgrove, 2005). By providing cell wall plasticity, expansins play an effective role in plant growth and development (Pien et al., 2001; Kuluev et al., 2012, 2013; Zhou et al., 2015), seed germination (Huang et al., 2000; Chen et al., 2001, (Morris et al., 2011; Yan et al., 2014), pollen tube penetration into the stigma (Valdivia et al., 2007), fruit ripening and softening (Civello et al., 1999), fibre development (Shimizu et al., 1997), and biotic and abiotic stress response (Han et al., 2015; Zhou et al., 2015; Zhang et al., 2018; Cosgrove, 2015; Jin, et al., 2020).

Cannabis sativa is a plant of economic value with many different uses such as textile, food, paper, bioplastic, insulation, biofuel, and pharmaceutical industry. Thus, *Cannabis sativa* is divided into two separate varieties for industrial (hem) and pharmaceutical (marijuana) uses. Textile products made from

the fashion industry is embracing hemp fibers because The fabric is breathable, lightweight, tough and durable (Islam et al., 2021). Additionally, the ability to grow hemp without the need for pesticides or herbicides is important for sustainability. In the use of hemp in all these areas, knowing the structural properties of the cell wall can enable the product to be developed in the desired direction. Cannabis marijuana is especially important for the production of THC and CBD for medical use.

In this study, genome-wide a characterization of the expansin genes in the cannabis genome was carried out to determine their structural properties and evolutionary relationships, chromosomal locations, the cisacting elements involved in the development, response to light, hormones, and environmental stress, and in silico gene expression.

2. Materials and Methods

The expansin protein sequences of A. thaliana (ATEXPAs, ATEXBs, ATEXLAs, ATEXLBs) derived from TAIR (https://www.arabidopsis.org/) were used as queries to identify expansin genes in Cannabis genome (ASM2916894v1, sativa NCBI https://www.ncbi.nlm.nih.gov/datasets/genom e/GCF 029168945.1/) using BLASTp with a (1e-5) E-value. All candidate expansin protein sequences were submitted to the HMMER v2.43 online program (https://www.ebi.ac.uk/Tools/hmmer/search/p hmmer) to identify conserved domains. The proteins were checked using DPBB_1 domain Lytic transglycolase (PF03330) and expansin C-terminal domain (pollen allergen domain, PF01357) at Pfam (http://pfam.xfam.org/) (Mistry et al., 2021). The signal peptide and the transmembrane domains were detected using the HMMER v2.43 online program.

The phylogenetic tree was drawn according to the Neighbor-joining method (Saitou and Nei, 1987) with 1000 replicates using MEGA11 (Tamura et al., 2021). ClustalW was performed to align amino acid sequences of expansin protein from *C. Sativa, A. thaliana*, and *Oryza sativa. O. sativa* and *A. thaliana*

expansin sequences were retrieved from TAIR (www.arabidopsis.org/), the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/), and the Expansin Central website (http://www.personal.psu.edu/fsl/ExpCentral/i ndex.html). MEME v5.5.2 (https://memesuite.org/meme/tools/meme) (Bailey et al., 2009) was utilized to find conserved 20 motifs. The gene structures were detected using the structure gene display server v2.0 (http://gsds.gao lab.org/). The chromosome mapping of the expansin genes on the cannabis chromosome was using the locations information of expansin genes derived NCBI (https://www.ncbi.nlm.nih.gov/). The syntenic relationships with A. thaliana were detected. The 2000-bp upstream of gene sequences retrieved from **NCBI** (https://www.ncbi.nlm.nih.gov/) were used as queries on the **PLANTCARE** (http://bioinformatics.psb.ugent.be/ webtools/plantcare/html/) (Lescot et al., 2002) to find cis-acting elements. Heatmaps and chromosome map were drawn using TBtools-II (Chen et al., 2023).

For enrichment analysis, functional classification of cannabis expansins was conducted with the DAVID Gene Functional Classification online tool (https://david.ncifcrf.gov/). The gene ontology terms were assumed as significantly enriched according to the P values ($P \le 0.05$).

3. Results and Discussion

A total of 29 expansin genes were found in C. sativa genome (ASM2916894v1). The expansin genes have been named and numbered according to species name, subfamily, and chromosomal expansion location (Table 1). Cannabis expansins were classified into four subfamilies, including eighteen CsEXPAs, five CsEXPBs, one CsEXLAs, and five CsEXLBs family members. All of them contained DPBB_1 domain and pollen allergen domain. The number of expansin genes detected in many species varied, regardless of genome size for example, 38 in tomato (Lu et al., 2016), 39 in potato (Li et al. 2020), 52 in tobacco (Ding et al., 2016), 58 in rice (Hemalatha et al., 2011), 75 in soybean (Feng et al., 2022) and 88 in maize (Zhang et al., 2014) and 93, 49 and 45 in the *G*. hirsutum, G. arboreum and G. raimondii (Lv et al., 2020), respectively.

The sequence length of cannabis expansin proteins varied from 245 aa (CsEXPA14) to 339 aa (CsEXPA18). The molecular weight ranged from 26.41 kDa (CsEXPA14) to 37.99 (CsEXPA18) kDa. The isoelectric point (pI) also varied from 4.82 to 10.12 (Table 1).

Sipahi

	Gene name	Protein ID	Molecular weights (kDa)	pI	Length (aa)
1	CsEXPA1	XP_030488120.1	26.98	8.37	254
2	CsEXPA2	XP_030490280.1	27.73	9.43	258
3	CsEXPA3	XP_030490239.1	28.83	10.12	257
4	CsEXPA4	XP_030506062.1	27.31	8.62	255
5	CsEXPA5	XP_030503724.1	27.95	9.63	258
6	CsEXPA6	XP_030492538.1	26.88	8.99	249
7	CsEXPA7	XP_030493381.2	31.78	9.38	283
8	CsEXPA8	XP_030502783.1	28.76	9.62	262
9	CsEXPA9	XP_030509858.1	26.91	9.39	251
10	CsEXPA10	XP_030479717.2	28.10	9.45	259
11	CsEXPA11	XP_060957840.1	28.14	9.63	259
12	CsEXPA12	XP_030481299.1	27.87	8.06	259
13	CsEXPA13	XP_030482760.1	26.44	9.32	248
14	CsEXPA14	XP_030483891.2	26.41	9.42	245
15	CsEXPA15	XP_030483935.1	27.73	8.4	250
16	CsEXPA16	XP_030500063.1	27.79	9.32	255
17	CsEXPA17	XP_030508707.1	27.51	9.59	257
18	CsEXPA18	XP_030486138.2	37.99	9.33	339
19	CsEXPB1	XP_030488027.1	29.26	9.02	272
20	CsEXPB2	XP_030502623.2	29.69	7.53	279
21	CsEXPB3	XP_030500448.2	29.11	6.4	278
22	CsEXPB4	XP_030501582.2	29.11	6.4	278
23	CsEXPB5	XP_030507862.2	29.02	8.9	267
24	CsEXPLA1	XP_030485441.1	28.12	8.6	261
25	CsEXPLB1	XP_030501721.2	29.33	9.21	261
26	CsEXPLB2	XP_030501803.2	29.30	9.15	261
27	CsEXPLB3	XP_030501763.1	29.37	9.21	262
28	CsEXPLB4	XP_030500425.2	27.90	6.07	253
29	CsEXPLB5	XP_030502993.2	28.04	4.82	255

Table 1. The characteristics of expansin proteins in C. sativa

The evolutionary relationships of cannabis expansins were shown in the phylogenetic tree (Fig. 1). The expansins from *Arabidopsis* (35) and rice (55) which have already been classified were included in this phylogenetic tree. The four groups of cannabis expansins (CsEXPA, CsEXPB, CsEXLA, and CsEXLB) were classified into four separate groups according to their similarity to homologous of *A. thaliana* and *O. sativa*. All expansins are divided into 10 subgroups. EXPA was the largest group, which had five subgroups with seventy-seven members, EXLA was the smallest group, including fifteen expansin members (Fig. 1). The fact that cannabis expansins were clustered in the same groups with Arabidopsis and rice expansin proteins indicates that expansin genes evolved before the monocot-dicot divergence.

Sipahi



Figure 1. Phylogenetic analysis of the expansin family members of C. sativa, A. thaliana, and O. sativa.

Most of cannabis expansin proteins were found in extracellular membrane, vacuole and

chloroplast (Fig. 2). Only CsEXPB2 was found in endoplasmic reticulum plasma.



Figure 2. The subcellular localization of cannabis expansin proteins.

The 29 *expansin* genes are distributed on nine chromosomes (Fig. 3). While *CsEXPLB* genes are located contiguously on

chromosome 5, *CsEXPA* genes are distributed on all chromosomes except chromosome 4.



Figure 3. Chromosomal locations of expansin genes in cannabis chromosomes

The synteny analysis revealed syntenic relationships between *C. sativa* and *A. thaliana* (Fig. 4). Thirteen expansin genes (*CsEXPA1*, -

2, -6, -8, -10, -14, -15, -16, -17, -18, *CsEXPB* 3, -5, *CsEXPLB1*) were found to be orthologous with *A. thaliana* expansins.



Figure 4. Syntenic relationship between *C. sativa* and *A. thaliana*. Red lines indicate syntenic lines while the grey color shows shared genomic blocks between the organisms.

To assess the divergence between the *expansin* genes, various conserved motifs and the pattern of their distribution between the genes were detected. Cannabis expansins have 6-9 different motifs, with the least motif being CsEXPLA1. Most of the cannabis *expansin* genes (23 genes) have three exons (Fig. 5). Only *CsEXPA5* contains two exons, while *CsEXPLA1* and *CsEXPLB5* contain five exons.

CsEXPA7, -11, and *CsEXPLB52* have the longest introns.

The cis-acting elements on the promoter of the cannabis expansin genes are shown in Fig. 6. A total of 69 cis-elements associated with development, hormones, light, and response to environmental stimuli have been identified in the promoters of expansin genes, indicating the multifunctional role of expansins. Twelve development-related cis-elements were discovered. AAGAA-motif (involved in the endosperm-specific negative expression), as-1 root-specific elements, and circadian cis-acting elements were especially found in many *expansin* gene promoters. The role of expansins in cell wall modification during root hair morphogenesis has been proven through various studies. Expression of the β -expansin *HvEXPB1* gene containing five root hair-specific cis-elements (RHEs) has been shown to be root hair specific and associated with root hair formation in barley (Won et al., 2010).

Also, the soybean expansin *GmEXPB2* gene was found to be involved in root hair elongation (Guo et al., 2011). Rice a β expansin gene (*OsEXPB2*) is found to have a role in rice root system architecture (Zou et al., 2015). Cho and Cosgrove (2000) indicated maximum expression of *AtEXP10* in the growing leaf and at the base of the pedicel and concluded that expansins have developmental roles of organ size, morphology, and abscission.



Figure 5. a) Phylogenetic relationships of the cannabis expansins. b) motifs found in the protein sequences of cannabis expansins. c) exon–intron organizations of cannabis expansion genes. exons and introns were shown by yellow boxes and black lines, respectively. twenty motifs were shown by different colored boxes.

Fifteen of environmental-stress-related cisacting elements were observed. These are LTR (low-temperature responsiveness), MBS (drought-inducibility), TC-rich repeats (stress responsiveness), ARE (anaerobic induction), DRE (dehydration, low-temp, and salt stresses), W box and WUN-motif (woundresponsive elements). The abscisic-acidresponsive cis-elements, including ABRE, ABRE2, ABRE3a, and ABRE4, were found in many expansin gene promoters. 23 expansin genes possessed the ethylene-responsive element (ERE). MeJA-responsive element (TGACG-motif, TATC-box, and CGTCAmotif), auxin-responsive (AuxRR-core, TGAand TGA-element), gibberellinbox, responsive element (GARE-motif), and salicylic-acid-responsive (TCA, TCA-

element) were detected in several *expansin* gene promotors. Finally, most cannabis

expansin genes possess Box 4, G-Box, and TCT-motif light-related cis-elements.



Figure 6. a) Development-related cis-acting elements, b) environmental stress response cis-elements, c) hormone response cis-elements, d) the light response cis-elements in cannabis expansin gene promoters.

10 and 8 cannabis expansins that are ortologous to *Arabidopsis* interacted with XTH23 (Probable xyloglucan endotransglucosylase/hydrolase protein 23) and XTH8 (probable xyloglucan endotransglucosylase/hydrolase protein 8; Catalyzes xyloglucan endohydrolysis), respectively (Figure 7). XTH23 and XTH8 cleave and religates xyloglucan polymers, thereby participating in the cell wall construction of growing tissues.



Figure 7. Protein-protein interaction of cannabis expansion

When the investigating data of the transcriptome changes of drought-stressed cannabis cultivar, compared to irrigated control cannabis cultivar by Gao et al. (2018), it was found that *CsEXPA11* (-3.35 fold), *CsEXPA13* (-1.5 fold), *CsEXPA14* (-2.32 fold), *CsEXPA3* (-1.91 fold) and *CsEXPA8* (-3.22 fold) genes were down-regulated. When trichomes and root tissue transcriptome data (Yeo et al. 2022) were compared, *CsEXPA13* showed down-regulation in trichomes, while *CsEXPA17* showed up-regulation in all

cultivars. GO enrichment analysis of cannabis expansins showed that the identified expansins were classified into two main categories: biological processes and cellular components (Figure 8A). In biological processes, most of the cannabis expansins were involved in cell wall organization or biogenesis. The enrichment analysis of this protein family based on cellular component showed that these proteins are mainly extracellular region and membrane (Figure 8B).



Figure 8. Biological process and cellular component of expansin in cannabis

4. Conclusion

Through the action of proteins such as wall is constantly expansins, the cell reshaped disassembled assembled, and throughout the life of the plant. This process is very important in the growth of the plant and its response to environmental factors. The effect of expansion on the cell wall is to provide plasticity. In this study, expansin proteins in cannabis plants were identified bioinformatically. This basic information is for future product improvement vital programs, especially in studies such as fiber development.

References

Chen, F., Dahal, P., Bradford, K.J., 2001. Two tomato expansin genes show divergent expression and localization in embryos during seed development and germination. *Plant Physiology*, 127(3): 928-936.

- Chen, C., Wu, Y., Li, J., Wang, X., Zeng, Z., Xu, J., Liu, Y., Feng, J., Chen, H., He, Y., Xia, R., 2023. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Molecular Plant* 16: 1733–1742.
- Cho, H.T., Cosgrove, D.J., 2000. Altered expression of expansin modulates leaf growth and pedicel abscission in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences*, 97(17): 9783-9788.
- Cosgrove, D.J., 2015. Plant expansins: diversity and interactions with plant cell walls. *Current Opinion in Plant Biology*, 25: 162-172.
- Civello, P.M., Powell, A.L., Sabehat, A., Bennett, A.B., 1999. An expansin gene expressed in ripening strawberry fruit. *Plant Physiology*, 121(4): 1273-1279.

- Ding, A., Marowa, P., Kong, Y., 2016. Genome-wide identification of the expansin gene family in tobacco (Nicotiana tabacum). *Molecular Genetics and Genomics*, 291(5): 1891-1907.
- Feng, X., Li, C., He, F., Xu, Y., Li, L., Wang, 2022. Х., Li, F., Genome-wide identification of Expansin genes in wild soybean (Glycine soja) and functional Expansin characterization of **B**1 (GsEXPB1) hair in soybean root. Journal International of Molecular Sciences, 23(10): 5407.
- Gao, C., Cheng, C., Zhao, L., Yu, Y., Tang, Q.,
 Xin, P., Zang, G., 2018. Genome-wide expression profiles of hemp (*Cannabis sativa* L.) in response to drought stress. *International Journal of Genomics*.
- Guo, W., Zhao, J., Li, X., Qin, L., Yan, X., Liao, H., 2011. A soybean β-expansin gene GmEXPB2 intrinsically involved in root system architecture responses to abiotic stresses. *The Plant Journal*, 66(3): 541-552.
- Han, Y., Chen, Y., Yin, S., Zhang, M., Wang, W., 2015. Over-expression of TaEXPB23, a wheat expansin gene, improves oxidative stress tolerance in transgenic tobacco plants. *Journal of Plant Physiology*, 173: 62-71.
- Hemalatha, N., Rajesh, M.K., Narayanan, N.K., 2011. Genome-wide analysis and identification of genes related to expansin gene family in indica rice. *International Journal of Bioinformatics Research and Applications*, 7(2): 162-167.
- Huang, J., Takano, T., Akita, S., 2000. Expression of α-expansin genes in young seedlings of rice (*Oryza sativa* L.). *Planta* 211(4): 467–473
- Islam, M.S., Ahmed, S., Azady, M.A.R., 2021. Sustainable technologies for textile production. In *Fundamentals of Natural Fibres and Textiles* (pp. 625-655). Woodhead Publishing
- Jin, K.M., Zhuo, R.Y., Xu, D., Wang, Y.J., Fan, H.J., Huang, B.Y., Qiao, G.R., 2020. Genome-wide identification of the expansin

gene family and its potential association with drought stress in moso bamboo. *International Journal of Molecular Sciences*, 21(24): 9491.

- Kuluev, B.R., Mikhaylova, E.V., Chemeris, A.V., 2013. Transfer of the ARGOS-LIKE and AtEXPA10 genes into non-transgenic forms of tobacco and the phenotypic effects of their constitutive expression. *Russian Journal of Genetics: Applied Research*, 3: 265-270.
- Kuluev, B.R., Knyazev, A.B., Lebedev, Y.P., Chemeris, A.V., 2012. Morphological and physiological characteristics of transgenic tobacco plants expressing expansin genes: AtEXP10 from Arabidopsis and PnEXPA1 from poplar. *Russian Journal of Plant Physiology*, 59: 97-104.
- Li, Y., Lin-Wang, K., Liu, Z., Allan, A.C., Qin, S., Zhang, J., Liu, Y., 2020. Genome-wide analysis and expression profiles of the StR2R3-MYB transcription factor superfamily in potato (*Solanum tuberosum* L.). *International Journal of Biological Macromolecules*, 148: 817-832.
- Lu, Y., Liu, L., Wang, X., Han, Z., Ouyang, B., Zhang, J., Li, H., 2016. Genome-wide identification and expression analysis of the expansin gene family in tomato. *Molecular Genetics and Genomics*, 291: 597-608.
- Lv, L.M., Zuo, D.Y., Wang, X.F., Cheng, H. L., Zhang, Y.P., Wang, Q.L., Ma, Z.Y., 2020. Genome-wide identification of the expansin gene family reveals that expansin genes are involved in fibre cell growth in cotton. *BMC Plant Biology*, 20: 1-13.
- Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G.A., Sonnhammer, E.L.L., Tosatto, S.C.E., Paladin, L., Raj, S., Richardson, L.J., Finn, R.D., Bateman, A., 2021. Pfam: The protein families database in 2021. *Nucleic Acids Research*, 49(D1): D412–D419.

- Morris, K., Linkies, A., Müller, K., Oracz, K., Wang, X., Lynn, J.R., Finch-Savage, W.E., 2011. Regulation of seed germination in the close Arabidopsis relative Lepidium sativum: a global tissue-specific transcript analysis. *Plant Physiology*, 155(4): 1851-1870.
- Pien, S., Wyrzykowska, J., McQueen-Mason, S., Smart, C., Fleming, A., 2001. Local expression of expansin induces the entire process of leaf development and modifies leaf shape. *Proceedings of the National Academy of Sciences*, 98(20): 11812-11817.
- Saitou, N., Nei, M., 1987. The neighborjoining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4): 406–425.
- Sampedro, J., Cosgrove, D.J., 2005. The expansin superfamily. *Genome Biology*, 6: 1-11.
- Shimizu, Y., Aotsuka, S., Hasegawa, O., Kawada, T., Sakuno, T., Sakai, F., Hayashi, T., 1997. Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fiber cells. *Plant Cell Physiology* 38(3): 375–378.
- Tamura, K., Stecher, G., Kumar, S., 2021 MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7): 3022–3027.
- Valdivia, E.R., Wu, Y., Li, L.C., Cosgrove, D.J., Stephenson, A.G., 2007. A group-1

grass pollen allergen influences the outcome of pollen competition in maize. *PLoS One*, 2(1): e154.

- Won, S.K., Choi, S.B., Kumari, S., Cho, M., Lee, S.H., Cho, H.T., 2010. Root hairspecific EXPANSIN B genes have been selected for Graminaceae root hairs. *Molecules and Cells*, 30: 369-376.
- Yeo, H.C., Reddy, V.A., Mun, B.G., Leong, S.H., Dhandapani, S., Rajani, S., Jang, I.C., 2022. Comparative transcriptome analysis reveals coordinated transcriptional regulation of central and secondary metabolism in the trichomes of cannabis cultivars. *International Journal of Molecular Sciences*, 23(15): 8310
- Zhang, J.F., Xu, Y.Q., Dong, J.M., Peng, L.N., Feng, X., Wang, X., Li, F.L., 2018. Genome-wide identification of wheat (*Triticum aestivum*) expansins and expansin expression analysis in cold-tolerant and cold-sensitive wheat cultivars. *PLoS One*, 13(3): e0195138.
- Zhou, S., Han, Y.Y., Chen, Y., Kong, X., Wang, W., 2015. The involvement of expansins in response to water stress during leaf development in wheat. *Journal of Plant Physiology*, 183: 64-74.
- Zou, H., Wenwen, Y., Zang, G., Kang, Z., Zhang, Z., Huang, J., Wang, G., 2015. OsEXPB2, a β-expansin gene, is involved in rice root system architecture. *Molecular Breeding*, 35: 1-14.

To CiteSipahi, H., 2024. Genome-Wide Characterization of Expansin Gene Family in
Cannabis sativa. ISPEC Journal of Agricultural Sciences, 8(1): 255-265.
DOI: https://doi.org/10.5281/zenodo.10864859.