Journal of Tropical Biodiversity and Biotechnology
Volume 10, Issue 04 (2025): jtbb15909
DOI: 10.22146/jtbb.15909

Research Article

Bioinformatic Characterization of the Mitogen-Activated Protein Kinase Genes in Wild *Arachis* Species with Expression Insights in *Arachis hypogaea*

Ha Duc Chu[1], Huyen Thi Thanh Tran[2], Le Thi Ngoc Quynh[3], Dong Huy Gioi[4], Hong Viet La[5], Man Thi Le[6], Phi Bang Cao[6]*

1)Faculty of Agricultural Technology, University of Engineering and Technology, Vietnam National University Hanoi, Xuan Thuy Road, Cau Giay Ward, Hanoi City 122300, Vietnam
2)Faculty of Biology, Hanoi National University of Education, Xuan Thuy Road, Cau Giay Ward, Hanoi City 122300, Vietnam
3)Department of Biotechnology, Thuyloi University, Tay Son Road, Kim Lien Ward, Hanoi City 122300, Vietnam
4)Faculty of Biotechnology, Vietnam National University of Agriculture, Gia Lam Ward, Hanoi City 122300, Vietnam
5)Institute of Research and Application, Hanoi Pedagogical University 2, Xuan Hoa Ward, Phu Tho Province 280000, Vietnam
6)Faculty of Natural Sciences, Hung Vuong University, Nong Trang Ward, Phu Tho Province 35000, Vietnam
* Corresponding author, email: phibang.cao@hvu.edu.vn

Keywords:
Evolutionary genomics
Mitogen-activated protein kinase
Peanut
Tissue-specific expression
Wild *Arachis*
Submitted:
20 March 2024
Accepted:
07 May 2025
Published:
06 October 2025
Editors:
Ardaning Nuriliani
Liya Audinah

ABSTRACT

Mitogen-activated protein kinases (MAPKs) are crucial signalling components involved in plant growth, development, and responses to environmental stimuli. While their roles are well established in model plants, comprehensive characterisation in *Arachis* species remains limited. This *in silico* study conducted a genome-wide identification and analysis of *MAPK* genes in cultivated peanut (*A. hypogaea*) and its two wild *Arachis* species, *A. duranensis*, and *A. ipaensis*. We identified 42 AhMAPK, 18 AraduMAPK, and 18 AraipMAPK proteins in *A. hypogaea*, *A. duranensis*, and *A. ipaensis*, respectively. These MAPK proteins exhibited diverse physicochemical properties and gene structures. We constructed a maximum likelihood-based phylogenetic tree, categorising the MAPK proteins in *Arachis* species, *Arabidopsis thaliana*, and *Medicago truncatula*, into five distinct groups. Gene structure analysis indicated substantial exon-intron variation, implying potential regulatory complexity and alternative splicing mechanisms. Transcriptome data analysis across multiple major organ and tissue types revealed differential expression patterns, with certain *AhMAPK* genes showing strong tissue-specific expression, particularly in leaves, roots, and reproductive organs. The inclusion of diploid progenitors provided insights into the evolutionary trajectory and functional conservation of *MAPK* genes in *Arachis* species. These findings contribute to a deeper understanding of MAPK-mediated signalling in peanuts and offer a genetic foundation for future studies aimed at improving stress resilience and crop performance. The identified *MAPK* genes present valuable targets for genetic engineering and molecular breeding programmes to enhance peanut productivity and adaptability to environmental stresses.

How to cite:
Chu, H.D. et al., 2025. Bioinformatic Characterization of the Mitogen-Activated Protein Kinase Genes in Wild *Arachis* Species with Expression Insights in *Arachis hypogaea*. *Journal of Tropical Biodiversity and Biotechnology*, 10(4), jtbb15909. doi: 10.22146/jtbb.15909

INTRODUCTION

The cultivated peanut (*Arachis hypogaea*) is a multi-functional crop of significant agricultural and economic importance, primarily cultivated in tropical and subtropical regions around the world (Toomer 2018). Originating from South America, the peanut has an interesting history of domestication and dispersal, making its way across the globe through Spanish and Portuguese explorers (Bertioli et al. 2020; Wadood et al. 2022). Peanuts play a crucial role in global food systems, serving as a vital source of protein, fat, and other nutrients (Hill 2002; Toomer 2018). They are utilised in various forms, including whole peanuts, peanut oil, and peanut butter, contributing to both human nutrition and livestock feed. Despite its widespread cultivation and utility, peanuts face various abiotic stresses, such as drought, salinity, and temperature extremes, which can significantly impact yield and quality (Li et al. 2014; Zheng et al. 2019). Understanding the molecular mechanisms underlying peanut growth, development, and stress adaptation is essential for improving crop resilience. Several regulatory pathways, including hormonal signalling networks and transcription factor-mediated gene expression, have been implicated in peanut development and stress responses. However, the specific molecular players governing these processes remain largely unexplored (Kosev & Vasileva 2019).

Mitogen-activated protein kinase (MAPK) proteins are pivotal components of signalling pathways in plants, playing crucial roles in regulating cellular responses to a variety of external stimuli (Avruch 2007; Zhang & Zhang 2022). Specifically, MAPKs are serine/threonine protein kinases that are organised into three-tiered cascades, including MAP kinase kinase kinases, MAP kinase kinases, and MAPKs (Group 2002). This categorised arrangement facilitates the transmission of signals from the cell surface to the nucleus, enabling an effective and coordinated cellular response. Of particular interest, MAPKs are involved in a wide array of plant processes such as growth, development, and reproduction, and are particularly critical in the adaptation and survival strategies of plants under stress conditions (Jagodzik et al. 2018; Lin et al. 2021; Niekerk et al. 2024). Recently, the MAPK families have been comprehensively identified and characterized in a large number of higher plant species, including rice (*Oryza sativa*) (Reyna & Yang 2006), Arabidopsis (*Arabidopsis thaliana*) (Andreasson & Ellis 2010), purple false brome (*Brachypodium distachyon*) (Chen et al. 2012), tomato (*Solanum lycopersicum*) (Kong et al. 2012), three legume species, including *Lotus japonicus*, *Medicago truncatula* and *Phaseolus vulgaris* (Neupane et al. 2013), maize (*Zea mays*) (Liu et al. 2013), grapevine (*Vitis vinifera*) (Wang et al. 2014), cucumber (*Cucumis sativus*) (Wang et al. 2015), chrysanthemum (*Chrysanthemum morifolium*) (Song et al. 2018), chickpea (*Cicer arietinum*) (Singh et al. 2018), cultivated strawberry (*Fragaria* × *ananassa*) (Li et al. 2022), lettuce (*Lactuca sativa*) (Wang et al. 2022), as well as banana (*Musa* spp.) (Fan et al. 2023), where they are known to regulate responses to abiotic and biotic stresses. Despite their established roles in other crops, there have been no recent reports on the MAPKs in peanuts, even though the genome of peanuts has been released (Clevenger et al. 2016; Zhuang et al. 2019). Furthermore, *A. hypogaea* is an allotetraploid species derived from the hybridisation of two wild diploid ancestors, *A. duranensis* (A-genome) and *A. ipaensis* (B-genome). These wild species serve as essential genetic resources for understanding the evolutionary development of *A. hypogaea*, as well as for identifying key genes associated with agronomic traits. Investigating the MAPK protein family across these three species allows for a comparative genomic approach to explore their conservation, divergence, and functional relevance in peanut biology.

This study aimed to conduct a computational analysis of MAPK proteins in three *Arachis* species, including *A. hypogaea*, *A. duranensis*, and *A. ipaensis* (García et al. 2021). We began by identifying and annotating all putative MAPK proteins within the most recent *A. hypogaea*, *A. duranensis*, and *A. ipaensis* genome

assemblies. Using bioinformatics tools, we analysed the physical and chemical properties of each identified MAPK protein. Further investigations included exploring the classification and gene structure of these proteins. Finally, we re-analysed the expression levels of genes encoding the MAPK proteins across various major organs and tissues of the peanut by accessing previous microarray datasets. The findings provide valuable insights into the potential functions of *MAPK* genes in peanut development. As no prior genome-wide studies on *MAPK* genes in peanuts exist, this research establishes a foundation for future functional studies and genetic improvement efforts.

MATERIALS AND METHODS

Screening of the mitogen-activated protein kinase proteins in *Arachis* species

To identify members of the *MAPK* gene families in *Arachis* spp., a two-step approach was adopted as previously described (Brasileiro et al. 2023; Chu et al. 2024). Initially, a gene family search was conducted in the PeanutBase database (Dash et al. 2016), utilising the well-characterised MAPK proteins from *A. thaliana* (Andreasson & Ellis 2010). This search targeted the annotated genomes of cultivated peanut, namely *A. hypogaea* (BioProject: PRJNA419393) (Bertioli et al. 2019) and two wild *Arachis* species, including *A. duranensis* (BioProject: PRJNA258023) and *A. ipaensis* (BioProject: PRJNA258025) (Bertioli et al. 2016). Only non-redundant *MAPK* genes were selected for further analysis.

Characterisation of the mitogen-activated protein kinase proteins in *Arachis* species

The molecular weight (mW) and common physicochemical properties of each MAPK protein sequences were predicted using the ExPASy Protparam web-based tool (Gasteiger et al. 2003, 2005), as previously reported (La et al. 2022a, 2022b; Le et al. 2022). Among them, isoelectric point (pI), grand average of hydropathicity (GRAVY), and aliphatic index (AI) values were analysed. To maintain consistency in the analysis, all tools and databases were employed with their default settings throughout the process.

Phylogenetic analysis of the mitogen-activated protein kinase proteins in *Arachis* species

The construction of the phylogenetic tree for the well-characterised MAPK members from *Arabidopsis* (Andreasson & Ellis 2010), *M. truncatula* (Neupane et al. 2013), and three *Arachis* species was carried out by first aligning the full-length protein sequences using the ClustalX program (Thompson et al. 1997; Thompson et al. 2002). Following the alignment, a Maximum Likelihood-based phylogenetic tree was generated using MEGA11 software (Tamura et al. 2021). A total of 1000 bootstrap replicates were used to explore the phylogenetic relationships among members of the MAPK families as previously described (Cao 2022; La et al. 2022a, 2022b; Le et al. 2022; Chu et al. 2024).

Exon/intron organisation of the mitogen-activated protein kinase proteins in *Arachis hypogaea*

To analyse the gene structure of genes encoding the MAPK in *A. hypogaea*, we examined the organisation of the exons and introns for each MAPK gene as following previous studies (Cao 2022; La et al. 2022a, 2022b; Le et al. 2022; Chu et al. 2024). Using the BioEdit software (Hall 1999), we first calculated the lengths of the genomic DNA sequence for each gene. These genomic DNA sequence and coding DNA sequence data were then input into the Gene Structure Display Server (Hu et al. 2015), which facilitated the construction and visualisation of the exon/intron structure for each *MAPK* gene.

Expression analysis of the mitogen-activated protein kinase proteins in *Arachis hypogaea*

To explore the expression profiles of the *MAPK* genes across various growth and development stages in *A. hypogaea* plants, we re-analysed recent transcriptome datasets from the GEO NCBI (Barrett et al. 2013) and the PeanutBase websites (Dash et al. 2016). Specifically, we focused on an expression atlas from PeanutBase (Dash et al. 2016), which has been catalogued under GEO accession number GSE71357 (Clevenger et al. 2016). This microarray dataset was re-analysed to validate the Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values for the peanut *MAPK* genes across ten major organs and tissues, including mainstem leaf, lateral stem leaf, seedling leaf, vegetative shoot tip, reproductive shoot tip, nodule, root, perianth, stamen, and pistil (Clevenger et al. 2016). We also accessed one recent dataset, namely GSE180915, that provided the transcriptomic profiles in lateral and terminal leaflets, petiole, and shoot apical meristem. The re-analysis was conducted using R-script, following the methods previously described (La et al. 2022a, 2022b; Le et al. 2022).

RESULTS AND DISCUSSION

Genome-wide identification of the mitogen-activated protein kinase proteins in *Arachis* species

In this study, genome-wide searches focused on identifying conserved MAPK domains were conducted across the genomes of three *Arachis* species, including *A. duranensis* (Bertioli et al. 2016), *A. ipaensis* (Bertioli et al. 2016) and *A. hypogaea* (Bertioli et al. 2019). After manually removing redundant sequences, only proteins that contained conserved MAPK domains were retained. The naming of each *MAPK* gene family was based on the chromosomal locations for each species. As a result, a total of 42 AhMAPK proteins in *A. hypogaea*, 18 AraduMAPK proteins in *A. duranensis* and 18 AraipMAPK proteins in *A. ipaensis* were comprehensively identified (Table 1 and 2).

Recently, the MAPK proteins have been extensively screened across various plant species, revealing significant diversity in their numbers and functions. In *Arabidopsis*, 20 MAPK proteins have been documented (Andreasson & Ellis 2010), while rice has reported 15 members (Reyna & Yang 2006). Other plant species have shown varied counts, such as chrysanthemum (11 genes) (Song et al. 2018) , grape (12 genes) (Wang et al. 2014), and cultivated strawberry (43 genes) (Li et al. 2022). In the genomes of *L. japonicus*, *M. truncatula* and *P. vulgaris*, it has been identified varying numbers of MAPK proteins (Neupane et al. 2013). Specifically, *L. japonicus* possesses 19 MAPK proteins, *M. truncatula* has 18 MAPK proteins, while *P. vulgaris* contains 15 MAPK proteins (Neupane et al. 2013). In cucumber, a total of 14 members of the MAPK family have been reported (Wang et al. 2015). According to the banana genome databases, *M. acuminata*, *M. balbisiana*, *M. itinerans*, *M. schizocarpa*, and *M. textilis* contained 21 (namely MaMPK1-21), 12 (namely MbMPK1-12), 18 (namely MiMPK1-18), 16 (namely MsMPK1-16), and 10 (namely MtMPK1-10) members of the MAPK family, respectively (Fan et al. 2023). This study extended the exploration to three *Arachis* species, identifying 42 AhMAPK, 18 AraduMAPK, and 18 AraipMAPK proteins, which were designated from AhMAPK01 to AhMAPK42, AraduMAPK01 to AraduMAPK18 and AraipMAPK01 to AraipMAPK18 proteins in *A. hypogaea*, *A. duranensis*, and *A. ipaensis*, respectively.

Characterisation of the general properties of the mitogen-activated protein kinase proteins in *Arachis* species

The MAPK protein lengths varied across the three *Arachis* species, reflecting structural diversity. In *A. hypogaea*, 42 AhMAPK proteins ranged from 125

Table 1. Summary of the mitogen-activated protein kinase proteins in *Arachis hypogaea*.

Gene name	Locus name	gDNA (bp)	PL (aa)	mW (kDa)	pI	GRAVY	AI
AhMAPK01	arahy.Tifrunner.gnm1.ann1.LP1FMS	4788	612	69.51	9.20	-0.46	81.62
AhMAPK02	arahy.Tifrunner.gnm1.ann1.E5L63Q	3998	578	65.73	9.00	-0.45	79.84
AhMAPK03	arahy.Tifrunner.gnm1.ann1.YDHQ7Y	3419	383	43.86	5.91	-0.28	92.40
AhMAPK04	arahy.Tifrunner.gnm1.ann1.RQ7QZ9	2657	371	42.57	5.60	-0.23	99.65
AhMAPK05	arahy.Tifrunner.gnm1.ann1.Z8RB6C	2512	370	42.68	4.97	-0.29	96.97
AhMAPK06	arahy.Tifrunner.gnm1.ann1.HLI3FT	2027	393	44.50	6.32	-0.36	86.11
AhMAPK07	arahy.Tifrunner.gnm1.ann1.80H1WW	3235	397	45.40	5.51	-0.33	88.26
AhMAPK08	arahy.Tifrunner.gnm1.ann1.2MP0U5	4078	608	69.10	7.01	-0.63	80.05
AhMAPK09	arahy.Tifrunner.gnm1.ann1.M5D7FW	2102	368	42.27	8.00	-0.17	99.10
AhMAPK10	arahy.Tifrunner.gnm1.ann1.VK23H5	6635	563	63.96	8.80	-0.46	78.67
AhMAPK11	arahy.Tifrunner.gnm1.ann1.BC5GM2	5094	376	43.33	5.91	-0.34	91.54
AhMAPK12	arahy.Tifrunner.gnm1.ann1.HE88YK	2028	380	43.49	6.36	-0.35	90.34
AhMAPK13	arahy.Tifrunner.gnm1.ann1.U4IX7Q	5192	482	55.28	6.50	-0.53	81.97
AhMAPK14	arahy.Tifrunner.gnm1.ann1.3P0VVA	5296	663	75.85	9.19	-0.41	81.76
AhMAPK15	arahy.Tifrunner.gnm1.ann1.KXH2SR	4980	600	68.00	9.16	-0.49	78.72
AhMAPK16	arahy.Tifrunner.gnm1.ann1.X8RB79	1807	372	42.64	6.14	-0.20	97.55
AhMAPK17	arahy.Tifrunner.gnm1.ann1.ZIU9VW	4793	612	69.50	9.19	-0.46	81.78
AhMAPK18	arahy.Tifrunner.gnm1.ann1.521JDX	4285	577	65.65	9.09	-0.48	76.95
AhMAPK19	arahy.Tifrunner.gnm1.ann1.WCN1YP	3048	383	43.92	5.91	-0.27	92.40
AhMAPK20	arahy.Tifrunner.gnm1.ann1.YJFK71	1391	291	32.50	8.64	-0.19	96.15
AhMAPK21	arahy.Tifrunner.gnm1.ann1.AF4FIK	2664	371	42.58	5.60	-0.24	99.11
AhMAPK22	arahy.Tifrunner.gnm1.ann1.FN4EG0	2267	369	42.57	4.97	-0.30	96.18
AhMAPK23	arahy.Tifrunner.gnm1.ann1.773MH2	1996	387	44.05	6.32	-0.37	86.43
AhMAPK24	arahy.Tifrunner.gnm1.ann1.C5DY5V	3247	397	45.40	5.51	-0.33	88.26
AhMAPK25	arahy.Tifrunner.gnm1.ann1.68AXRD	4095	606	68.88	7.01	-0.63	79.98
AhMAPK26	arahy.Tifrunner.gnm1.ann1.68T27Q	1710	368	42.27	8.00	-0.17	99.10
AhMAPK27	arahy.Tifrunner.gnm1.ann1.9G5FU0	7429	563	63.97	8.80	-0.45	79.70
AhMAPK28	arahy.Tifrunner.gnm1.ann1.L410JY	4675	376	43.33	5.91	-0.34	91.54
AhMAPK29	arahy.Tifrunner.gnm1.ann1.CC6N12	2446	380	43.50	6.28	-0.35	89.84
AhMAPK30	arahy.Tifrunner.gnm1.ann1.7K4NIS	4577	482	55.31	6.50	-0.53	81.97
AhMAPK31	arahy.Tifrunner.gnm1.ann1.IN5DGA	746	125	14.02	10.00	-0.38	92.16
AhMAPK32	arahy.Tifrunner.gnm1.ann1.BPF1MW	6451	426	48.31	5.60	-0.36	81.06
AhMAPK33	arahy.Tifrunner.gnm1.ann1.Q98R2Z	5179	667	76.02	9.23	-0.40	81.14
AhMAPK34	arahy.Tifrunner.gnm1.ann1.X0F9HV	1760	372	42.59	6.14	-0.19	98.36
AhMAPK35	arahy.Tifrunner.gnm1.ann1.8F3XE1	4447	600	68.04	9.18	-0.49	78.72
AhMAPK36	arahy.Tifrunner.gnm1.ann1.4D26T3	2084	355	40.87	9.51	-0.11	101.07
AhMAPK37	arahy.Tifrunner.gnm1.ann1.RFCR2Y	6491	550	62.18	9.61	-0.316	78.73
AhMAPK38	arahy.Tifrunner.gnm1.ann1.1E12MN	7483	439	50.11	5.85	-0.267	86.36
AhMAPK39	arahy.Tifrunner.gnm1.ann1.C4CQD3	6491	550	62.18	9.61	-0.316	78.73
AhMAPK40	arahy.Tifrunner.gnm1.ann1.ZSL8AZ	7456	438	49.92	5.93	-0.269	85.66
AhMAPK41	arahy.Tifrunner.gnm1.ann1.UDY993	1248	415	47.35	9.04	-0.381	80.58
AhMAPK42	arahy.Tifrunner.gnm1.ann1.QN2FLP	1218	405	46.30	8.97	-0.34	82.57

Note: gDNA: gene size (bp), PL: protein length (amino acid residues), MW: Molecular weight (kDa), pI: Isoelectric point, GRAVY: Grand average of hydropathicity, AI: Aliphatic index.

(AhMAPK31) to 667 (AhMAPK33) amino acids, averaging 383 amino acids. In *A. duranensis*, 18 AraduMAPK proteins ranged from 360 (AraduMAPK08) to 669 (AraduMAPK14) amino acids, with an average of 415 amino acids. Similarly, *A. ipaensis* contained 18 AraipMAPK proteins, ranging from 368 (AraipMAPK09) to 667 (AraipMAPK14) amino acids, also averaging 415 amino acids. The mW values of these MAPK proteins varied from 14.02 (AhMAPK31) to 76.02 (AhMAPK33) kDa (in the AhMAPK family from *A. hypogaea*), 41.02 (AraduMAPK08) and 76.41 (AraduMAPK14) kDa (in the AraduMAPK family from *A. duranensis*) and 42.27 (AraipMAPK09) and 76.02

Table 2. Summary of the mitogen-activated protein kinase proteins in *A. duranensis* and *A. ipaensis*.

Gene name	Locus name	gDNA (bp)	PL (aa)	mW (kDa)	pI	GRAVY	AI
AraduMAPK01	Aradu.271A7	3072	600	68.03	9.17	-0.41	83.25
AraduMAPK02	Aradu.B41M6	3229	578	65.67	9.00	-0.45	79.84
AraduMAPK03	Aradu.CGX1Z	3043	371	42.56	5.86	-0.30	91.46
AraduMAPK04	Aradu.IP5K7	2656	371	42.57	5.60	-0.23	99.65
AraduMAPK05	Aradu.RVT4Y	2513	369	42.54	4.97	-0.29	97.24
AraduMAPK06	Aradu.J97W2	1989	371	42.74	6.15	-0.31	90.94
AraduMAPK07	Aradu.A68MB	3576	607	68.97	7.01	-0.63	80.18
AraduMAPK08	Aradu.44GS0	3129	360	41.02	5.72	-0.34	88.94
AraduMAPK09	Aradu.Q0IL1	1044	368	42.29	8.00	-0.17	99.10
AraduMAPK10	Aradu.WE8JU	4030	563	63.96	8.80	-0.46	78.67
AraduMAPK11	Aradu.NJQ5G	4433	383	44.00	5.54	-0.28	92.40
AraduMAPK12	Aradu.A7MKH	2040	384	43.74	6.36	-0.34	89.92
AraduMAPK13	Aradu.H6YZR	4783	500	57.42	6.94	-0.52	80.56
AraduMAPK14	Aradu.E1ZU5	5072	669	76.41	9.03	-0.40	81.48
AraduMAPK15	Aradu.93Y1F	4763	600	68.00	9.16	-0.49	78.72
AraduMAPK16	Aradu.49RUG	2385	372	42.64	6.14	-0.20	97.55
AraduMAPK17	Aradu.4FA0W	4435	417	47.50	9.49	-0.45	75.30
AraduMAPK18	Aradu.IX021	7290	453	51.61	5.79	-0.21	87.55
AraipMAPK01	Araip.RV49X	4100	600	68.04	9.16	-0.42	83.25
AraipMAPK02	Araip.Y7IQI	3278	561	63.76	8.98	-0.50	77.06
AraipMAPK03	Araip.32FLV	3048	371	42.56	5.80	-0.30	91.46
AraipMAPK04	Araip.IL3I2	2664	371	42.58	5.60	-0.24	99.11
AraipMAPK05	Araip.8K6RH	2267	369	42.57	4.97	-0.30	96.18
AraipMAPK06	Araip.WRI31	1616	387	44.05	6.32	-0.37	86.43
AraipMAPK07	Araip.74VLD	5603	385	44.34	5.65	-0.34	89.22
AraipMAPK08	Araip.LMV71	3565	615	69.82	7.30	-0.61	80.55
AraipMAPK09	Araip.AT3RC	1710	368	42.27	8.00	-0.17	99.10
AraipMAPK10	Araip.96FUL	4009	563	63.97	8.80	-0.45	79.70
AraipMAPK11	Araip.AH33F	4211	376	43.33	5.91	-0.34	91.54
AraipMAPK12	Araip.6I993	2098	380	43.50	6.28	-0.35	89.84
AraipMAPK13	Araip.TJ3I8	4577	482	55.30	6.50	-0.53	81.97
AraipMAPK14	Araip.P1N43	4920	667	76.02	9.23	-0.40	81.14
AraipMAPK15	Araip.CL071	1760	372	42.60	6.14	-0.19	98.36
AraipMAPK16	Araip.U4Z8Q	4532	600	68.04	9.18	-0.49	78.72
AraipMAPK17	Araip.T0473	4135	397	44.70	9.56	-0.45	72.75
AraipMAPK18	Araip.Q5AHE	1218	392	44.72	9.20	-0.40	79.85

Note: gDNA: gene size (bp), PL: protein length (amino acid residues), MW: Molecular weight (kDa), pI: Isoelectric point, GRAVY: Grand average of hydropathicity, AI: Aliphatic index.

(AraipMAPK14) kDa (in the AraipMAPK family from *A. ipaensis*). Next, the pI values of the AhMAPK proteins in *A. hypogaea* spanned from 4.97 (AhMAPK22 and AhMAPK05) to 10.00 (AhMAPK31), while these scores of the AraduMAPK and AraipMAPK proteins in *A. duranensis* and *A. ipaensis* ranged from 4.97 (AraduMAPK05) to 9.49 (AraduMAPK17) and 4.97 (AraipMAPK05) to 9.56 (AraipMAPK17), respectively. Interestingly, we found that the GRAVY values of whole MAPK proteins in three *Arachis* species were minus, ranging from -0.48 (AhMAPK18) to -0.11 (AhMAPK36) in *A. hypogaea*, -0.63 (AraduMAPK07) to -0.17 (AraduMAPK09) in *A. duranensis*

and -0.61 (AraipMAPK08) to -0.17 (AraipMAPK09) in *A. ipaensis*. This finding strongly indicated that the MAPK proteins in these *Arachis* species were hydrophilic. Additionally, the AI values of the MAPK proteins in *A. hypogaea* ranged from 76.95 (AhMAPK18) to 101.07 (AhMAPK36), while the AI scores of the MAPK proteins in *A. duranensis* and *A. ipaensis* ranged from 75.30 (AraduMAPK17) to 99.65 (AraduMAPK04) and 72.75 (AraipMAPK17) to 99.11 (AraipMAPK04), respectively. Detailed data regarding the physicochemical properties of the MAPK proteins across the three *Arachis* species were comprehensively presented in Table 1 and 2.

The characteristics of the MAPK proteins in terms of their molecular profiles were thoroughly analysed in various higher plant species, revealing a range in their physicochemical properties. Particularly, in *Musa* species, the number of amino acids in these MAPK proteins ranged from 366 to 1155 residues (Fan et al. 2023). The mW of the MAPK proteins spanned from 41.68 to 130.81 kDa, while the theoretical pI ranged between 5.34 and 9.73, indicating diverse charge profiles across different MAPKs (Fan et al. 2023). The AI scores showed a range from 78.41 to 99.1, which typically suggests how thermostable the protein is, whereas the GRAVY results indicated that these MAPK proteins in *Musa* species were hydrophilic, further underscoring their compatibility with aqueous cellular environments (Fan et al. 2023). Additionally, the LsMAPK proteins in lettuce exhibited a range in length from 284 amino acids for LsMAPK4-3 to 782 amino acids for LsMAPK16-4 (Wang et al. 2022). Correspondingly, their mW values ranged from 33.00 kDa to 89.70 kDa (Wang et al. 2022). The theoretical pI values of these MAPK proteins also displayed a wide scope, extending from 4.74 for LsMAPK4-3 up to 9.08 for LsMAPK16, indicating significant variations in their charge characteristics across different molecular forms (Wang et al. 2022). These findings indicated that while MAPK proteins in *Arachis* share common hydrophilic characteristics with other plant species, they exhibited a more conserved molecular size and weight distribution, which may be linked to their functional roles in peanut signalling pathways. This suggested that MAPK proteins in *Arachis* species may have evolved under different selective pressures compared to species like *Musa*, potentially due to variations in developmental processes.

Phylogenetic analysis of the mitogen-activated protein kinase proteins in *Arachis* species

To investigate the relationships among the MAPK families in *A. hypogaea*, *A. duranensis*, and *A. ipaensis*, a Maximum Likelihood-based phylogenetic tree was constructed using MEGA software. *A. thaliana* and *M. truncatula* were selected as reference species due to their well-characterized MAPK families. *Arabidopsis* serves as a model dicot plant with a fully annotated genome (Andreasson & Ellis 2010), while *M. truncatula* is known as a legume species closely related to *Arachis* (Neupane et al. 2013). As provided in Figure 1, the phylogenetic tree was classified into five different groups, including groups A, B, C, D, and E. Particularly, group A contained four members of the AhMAPK family (AhMAPK04, AhMAPK07, AhMAPK21, and AhMAPK24), two members of the AraduMAPK family (AraduMAPK04 and AraduMAPK08) and two members of the AraipMAPK family (AraipMAPK04 and AraipMAPK07). Group B had ten, five, and five members of the AhMAPK, AraduMAPK, and AraipMAPK families in *A. hypogaea*, *A. duranensis*, and *A. ipaensis*, respectively, while group C exhibited four AhMAPK proteins (AhMAPK09, AhMAPK16, AhMAPK26, and AhMAPK34), two AraduMAPK proteins (AraduMAPK09 and AraduMAPK16) and two AraipMAPK proteins (AraipMAPK09 and AraipMAPK15). Group D shared the highest members of the MAPK families in three *Arachis* species, including 18 AhMAPK proteins, seven AraduMAPK proteins and seven AraipMAPK

proteins. Group E contained six AhMAPK proteins, two AraduMAPK proteins (AraduMAPK17 and AraduMAPK18) and two AraipMAPK proteins (AraipMAPK17 and AraipMAPK18).

Previously, the phylogenetic relationships and classification of the MAPK families in plant species have been comprehensively analysed. For example, a phylogenetic tree using the full-length amino acid sequences of 20 AtMAPK proteins from Arabidopsis, 12 FvMAPK proteins from *Fragaria vesca*, and 43 FaMAPK proteins from cultivated strawberry was constructed (Li et al. 2022). The analysis revealed that all MAPK proteins from Arabidopsis, *F. vesca*, and cultivated strawberry were divided into four distinct cluster groups, namely groups A, B, C, and D (Li et al. 2022). Another phylogenetic tree constructed for the MAPK protein family across Arabidopsis, *Helianthus annuus*, and chrysanthemum revealed a categorization into four distinct groups, including groups A, B, C, and D (Song et al. 2018). Next, phylogenetic analysis of the MAPK families in Arabidopsis and five *Musa* species, including *M. acuminata*, *M. balbisiana*, *M. itinerans*, *M. schizocarpa*, and *M. textilis*, clearly revealed that these MAPK proteins could be divided into groups A, B, C, and D based on phylogenetic relationships (Fan et al. 2023). In this study, the phylogenetic classification of MAPK proteins in *Arachis* species also identified distinct groupings, consistent with those observed in other plants. However, while the core MAPK classification aligns with previous studies, some species-specific variations were evident, particularly in *A. hypogaea*, which has undergone genome expansion due to polyploidization. The grouping of *Arachis* MAPKs alongside their counterparts from *A. thaliana* and *M. truncatula* suggested functional conservation within legume species, supporting their roles in signaling pathways associated with biological processes.

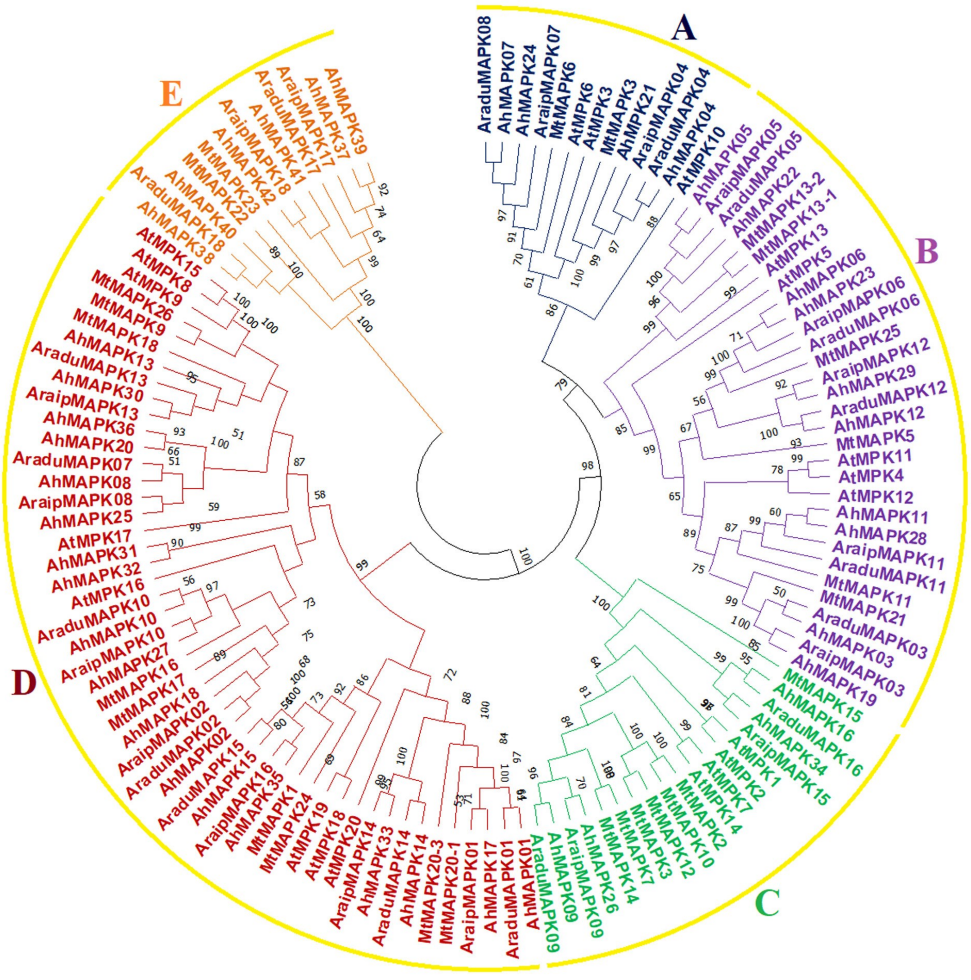

Figure 1. Classification of the MAPK families in *Arabidopsis thaliana*, *Medicago truncatula*, *Arachis hypogaea*, *A. duranensis*, and *A. ipaensis*.

Gene structure of the mitogen-activated protein kinase proteins in *Arachis* species

To examine the gene structure of the *MAPK* genes in *Arachis* species, the BioEdit software and Gene Structure Display Server (Hu et al. 2015) were applied to analyse the gene size and number of exons. As provided in Table 1, the genomic DNA sequences of the *AhMAPK* genes in *A. hypogaea* ranged from 746 (*AhMAPK31*) to 7483 (*AhMAPK38*) bp. Meanwhile, the gene sizes of the *AraduMAPK* and *AraipMAPK* genes in *A. duranensis* and *A. ipaensis* varied from 1044 (*AraduMAPK09*) to 7290 (*AraduMAPK18*) bp and 1218 (*AraipMAPK18*) to 5603 (*AraipMPK07*) bp, respectively (Tables 2).

Of particular interest, we focused on the exon/intron organization of the *AhMAPK* genes in *A. hypogaea* (Figure 2). Our analysis indicated that these *AhMAPK* genes displayed notable structural variations. For instance, two *AhMAPK* genes, including *AhMAPK41* and *AhMAPK42* each comprised a single exon, indicating a simpler gene structure. Conversely, the remaining 40 out of 42 *AhMAPK* genes had a variable number of exons, ranging from two to 19 exons. Four *AhMAPK* genes, including *AhMAPK09*, *AhMAPK16*, *AhMAPK26*, and *AhMAPK34* each consisting of two exons, while *AhMAPK31* and *AhMAPK20* showed an organization of four and five exons, respectively. Interestingly, 13 (out of 42) *AhMAPK* genes contained six exons. Only *AhMAPK36* had eight exons and two *AhMAPK* genes, namely *AhMAPK07* and *AhMAPK32* had nine exons, while seven and seven *AhMAPK* genes exhibited 10 and 11 exons, respectively. Finally, two (*AhMAPK38* and *AhMAPK40*) and two (*AhMAPK37* and *AhMAPK39*) genes contained 16 and 19 exons.

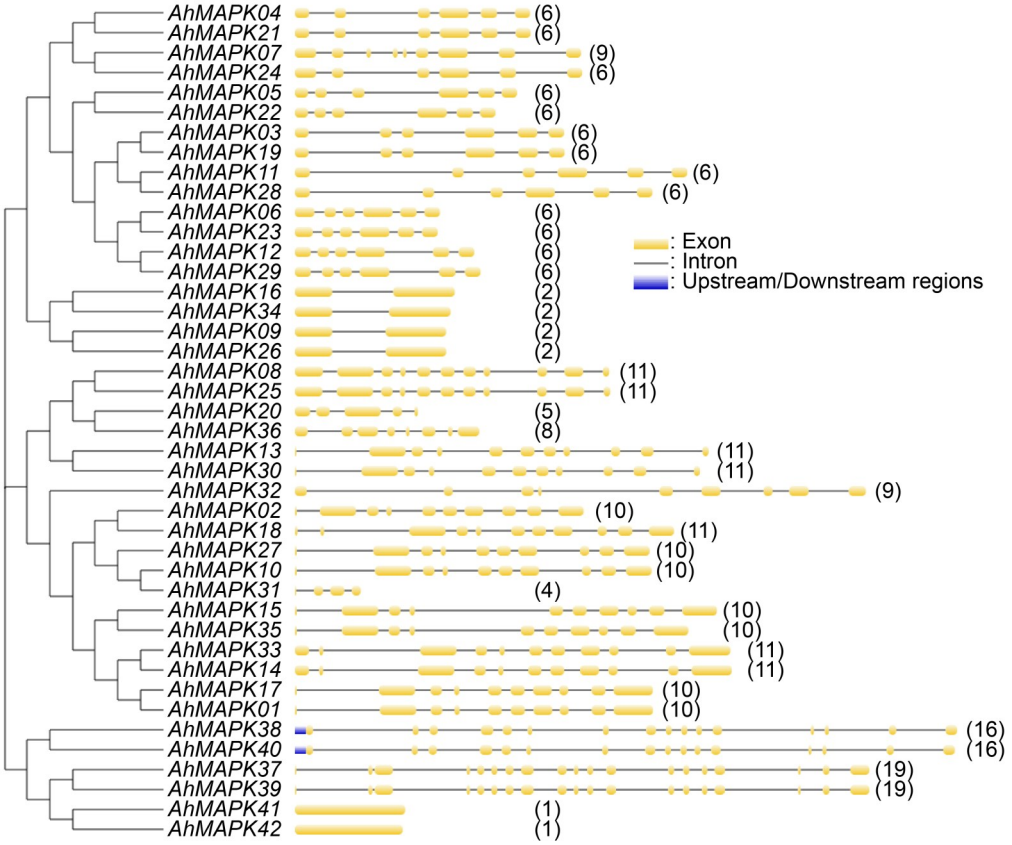

Figure 2. Gene structure of the MAPK family in *Arachis hypogaea*.

Previously, the gene structures encoding the MAPK families in plant species have been investigated. For example, it was discovered that all members of the MAPK family in purple false brome possess introns ranging from three to eleven, indicating a significant variation in the gene structure across this protein family (Chen et al. 2012). The number of introns within the

CsMAPK gene family in cucumber ranged from one to 10 (Wang et al. 2015), this finding was also confirmed in the *MAPK* gene families in tomato (Kong et al. 2012), grapevine (Wang et al. 2014), cucumber (Wang et al. 2015), chickpea (Singh et al. 2018), and lettuce (Wang et al. 2022). These variations suggest species-specific adaptations in *MAPK* gene structure, which may influence their functional roles in different plant lineages. In *A. hypogaea*, the exon-intron organization of *MAPK* genes also exhibits diversity, particularly among the 18 *AhMAPK* genes in Cluster D, which represents the largest MAPK subgroup. These genes showed a range of exon-intron structures, with some possessing multiple exons, indicating potential functional complexity. The presence of genes with higher intron numbers suggested possible regulation through alternative splicing. This structural variability aligns with the observed expression patterns of Cluster D *MAPK* genes. Such findings highlight the evolutionary significance of exon-intron organization in *A. hypogaea*, as the structural flexibility of these genes may contribute to their regulatory complexity. Understanding these structural characteristics provides essential insights into how *MAPK* genes in peanut function in developmental and environmental adaptation processes.

Expression patterns of the mitogen-activated protein kinase proteins in *Arachis hypogaea*

To explore the differential expression patterns of the *AhMAPK* genes in peanut, we analysed transcriptomic data for all 42 members from this gene family across 10 distinct organ and tissue samples (Figure 3). Our findings revealed a subset (17 out of 42) of the *AhMAPK* genes displaying notably lower expression levels in all tested tissues. Interestingly, a majority of the *AhMAPK* genes (25 out of 42) demonstrated significant expression, marked by high FPKM values, in at least one major organ or tissue. Notably, four *AhMAPK* genes, including *AhMAPK01*, *AhMAPK10*, *AhMAPK17*, and *AhMAPK27* exhibited the highest FPKM values in lateral stem leaf and mainstem leaf tissues, suggesting a specialized expression of these genes in these particular organs. Additionally, all *AhMAPK* genes presented lower expression levels in vegetative shoot tips and stamens, indicating a potential tissue-specific regulation of these *AhMAPK* genes. A total of 19 (out of 42) *AhMAPK* genes were exclusively expressed in root and/or nodule tissues, while five *AhMAPK* genes, particularly *AhMAPK01*, *AhMAPK10*, *AhMAPK15*, *AhMAPK17*, and *AhMAPK27*, were highly expressed in pistil tissues. Taken together, these findings suggest that the *AhMAPK* genes displayed a diverse range of expression patterns, underscoring the complexity of their role in peanut tissue specificity.

According to the GSE180915 dataset, the expression profiles of *AhMAPK* genes across four different organs, including lateral leaflets, petioles, shoot apical meristems, and terminal leaflets, indicated varying levels of tissue-specific expression (Figure 4). Several *AhMAPK* genes exhibited high expression across all four organs, suggesting their broad functional roles in peanut development. Notably, *AhMAPK37* and *AhMAPK39* displayed the highest expression levels in all tissues, particularly in terminal leaflets and lateral leaflets, indicating a potential role in leaf expansion and development. Similarly, *AhMPK14* showed strong expression in all organs, with the highest level in shoot apical meristem, suggesting its involvement in meristematic activity and growth regulation. The remaining *AhMAPK* genes showed minimal or negligible expression across all tissues.

Previously, the *MAPK* genes have been demonstrated to play a pivotal role in various developmental processes, with their tissue-specific expression patterns. Thus, the analysis of these patterns revealed crucial insights into the functional specialization of individual *MAPK* genes (Taj et al. 2010). For instance, higher expression of *MAPK* genes in root tissues may correlate with

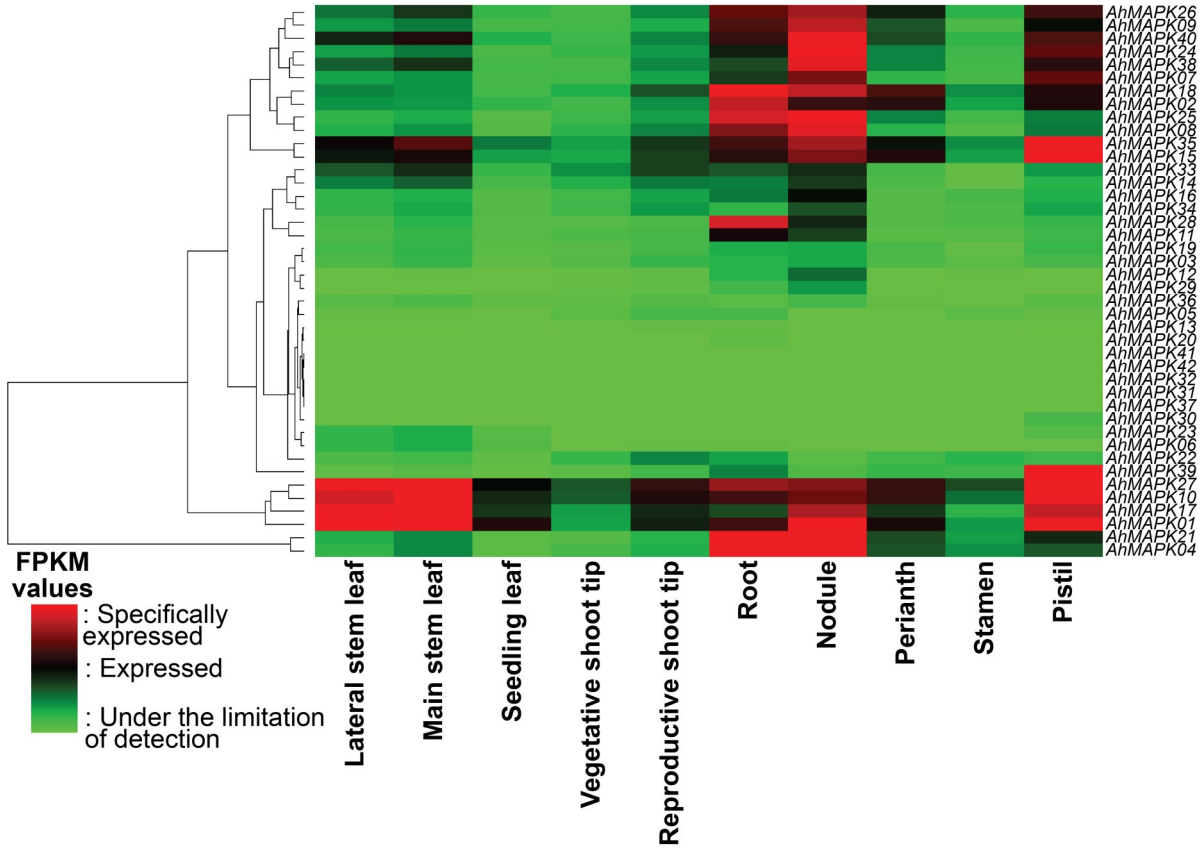

Figure 3. Expression heatmap of *AhMAPK* genes across ten major organs of *Arachis hypogaea*, including vegetative and reproductive tissues. The expression patterns of *AhMAPK* genes were explored across a comprehensive range of vegetative (e.g., lateral stem leaf, main stem leaf, seedling leaf, vegetative shoot tip, root, nodule) and reproductive organs (e.g., reproductive shoot tip, perianth, stamen, pistil) in *A. hypogaea*.

enhanced root growth or stress responses, such as those triggered by water deficiency (Opdenakker et al. 2012; Majeed et al. 2022). Similarly, the specific expression of other *MAPK* genes in the leaves and/or flowers could be linked to specific roles in photosynthesis regulation or reproductive development (Jonwal et al. 2023). Furthermore, *OsMPK3* in rice enhances drought tolerance by modulating stomatal closure (Reyna & Yang 2006), while *AtMPK6* in *Arabidopsis* plays a key role in pathogen defence by activating stress-responsive genes (Andreasson & Ellis 2010). In this study, the expression analysis revealed that several *MAPK* genes exhibited tissue-specific expression, with high transcript levels in leaves, roots, and reproductive organs. Notably, *AhMAPK10*, *AhMAPK17*, and *AhMAPK27* demonstrated high expression in lateral stem leaf and mainstem leaf tissues, indicating their potential involvement in leaf development and photosynthetic regulation. In contrast, several *MAPK* genes were more prominently expressed in root tissues, suggesting roles in root architecture or stress adaptation. Such differential expression not only underscores the adaptability of plants to their environment by modulating physiological and metabolic processes in a tissue-specific manner but also highlights the potential of these *MAPK* genes as targets for genetic engineering aimed at improving plant resilience and productivity (Taj et al. 2010). Therefore, understanding MAPK functions in peanuts could provide valuable insights into improving stress resilience and yield stability in peanut cultivation. Potential applications include using MAPK-associated genes as markers in breeding programs to develop drought-resistant varieties or leveraging biotechnological approaches to enhance peanut defense mechanisms against fungal pathogens such as *Aspergillus* species, which cause aflatoxin contamination. Future studies should focus on functional validation of *MAPK* genes in peanuts to explore their full potential in agricultural applications.

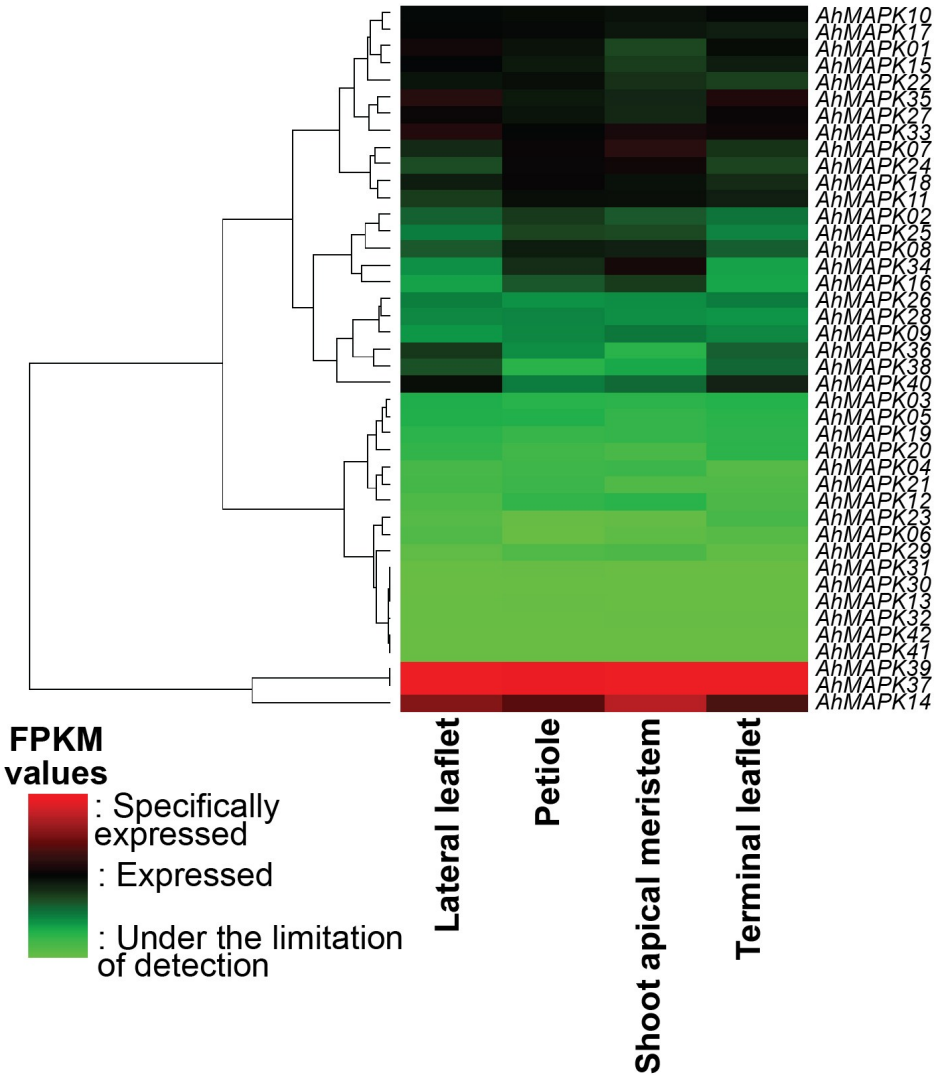

Figure 4. Tissue-specific expression profiles of *AhMAPK* genes in leaf-related structures of *Arachis hypogaea*. The expression levels of *AhMAPK* genes were explored in four aerial organs, including lateral leaflet, terminal leaflet, petiole, and shoot apical meristem.

CONCLUSION

This study presents a comprehensive genome-wide analysis of *MAPK* genes in *Arachis hypogaea*, *A. duranensis*, and *A. ipaensis*. A total of 42, 18, and 18 MAPK genes were identified in these species, respectively, and their molecular characteristics, gene structures, and evolutionary relationships were examined. The phylogenetic classification grouped these genes into five distinct clusters, revealing both conserved and divergent features among species. Expression analysis demonstrated tissue-specific patterns, with several *MAPK* genes showing high transcript levels in leaf and reproductive tissues, suggesting their roles in developmental processes and stress responses. Specifically, *AhMAPK10*, *AhMAPK17*, and *AhMAPK27* displayed strong expression in lateral and mainstem leaves, indicating their potential function in photosynthetic regulation, while *AhMAPK02*, *AhMAPK04*, *AhMAPK18*, *AhMAPK21*, *AhMAPK25*, and *AhMAPK28* were predominantly expressed in roots, likely contributing to root development and environmental adaptation. The findings of this study serve as a valuable genomic resource for understanding MAPK-mediated signalling in peanuts and provide a foundation for future applications in crop improvement. Further functional validation through transcriptomic and gene expression studies will be essential to explore the potential of these *MAPK* genes in enhancing stress tolerance and agricultural productivity.

AUTHORS CONTRIBUTION

H.D.C. contributed to the research design, data collection and analysis, and preparation of the first draft of the manuscript. H.T.T.T. contributed to data collection. L.T.Q.N. and D.H.G. contributed to data collection and analysis. H.V.L. contributed to the research design, data collection and analysis. M.T.L. contributed to data collection. P.B.C. contributed to the research design, data collection and analysis, preparation and editing of the manuscript, and supervision of the entire process.

ACKNOWLEDGMENTS

This work was fully funded by Vietnam National University Hanoi under the project number QG.25.61.

CONFLICT OF INTEREST

The authors declare no conflict of interest regarding the research or the research funding.

REFERENCES

Andreasson, E. & Ellis, B., 2010. Convergence and specificity in the Arabidopsis MAPK nexus. *Trends in Plant Science*, 15(2), pp.106-113. doi: 10.1016/j.tplants.2009.12.001

Avruch, J., 2007. MAP kinase pathways: the first twenty years. *Biochimica et biophysica acta*, 1773(8), pp.1150-1160. doi: 10.1016/j.bbamcr.2006.11.006

Barrett, T. et al., 2013. NCBI GEO: archive for functional genomics data sets - update. *Nucleic Acid Research*, 41(Database issue), pp.D991-995. doi: 10.1093/nar/gks1193

Bertioli, D.J. et al., 2016. The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nature Genetics*, 48(4), pp.438-446. doi: 10.1038/ng.3517

Bertioli, D.J. et al., 2019. The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. *Nature Genetics*, 51(5), pp.877-884. doi: 10.1038/s41588-019-0405-z

Bertioli, D.J. et al., 2020. Evaluating two different models of peanut's origin. *Nature Genetics*, 52(6), pp.557-559. doi:10.1038/s41588-020-0626-1

Brasileiro, A.C.M. et al., 2023. The stilbene synthase family in *Arachis*: A genome-wide study and functional characterization in response to stress. *Genes (Basel)*, 14(12), 2181. doi: 10.3390/genes14122181

Cao, P.B., 2022. *In silico* structural, evolutionary, and expression analysis of small heat shock protein (shsp) encoding genes in cocoa (*Theobroma cacao* L.). *The Journal of Animal and Plant Sciences*, 32(5), pp.1394-1402. doi: 10.36899/JAPS.2022.5.0546

Chen, L. et al., 2012. Genome-wide identification and analysis of MAPK and MAPKK gene families in *Brachypodium distachyon*. *PLoS One*, 7(10), e46744. doi:10.1371/journal.pone.0046744

Chu, H.D. et al., 2024. Identification of two enzymes for trehalose synthesis and their potential function in growth and development in peanut (*Arachis hypogaea*). *Journal of Tropical Life Science*, 14(1), pp.83-94. doi: 10.11594/jtls.14.01.10

Clevenger, J. et al., 2016. A developmental transcriptome map for allotetraploid *Arachis hypogaea*. *Frontiers in Plant Science*, 7, 1446. doi: 10.3389/fpls.2016.01446

Dash, S. et al., 2016. PeanutBase and other bioinformatic resources for peanut. In *Peanuts: Genetics, processing, and utilization*. AOCS Press, pp.241-252. doi: 10.1016/B978-1-63067-038-2.00008-3

Fan, Z. et al., 2023. Genome-wide identification of the MPK gene family and expression analysis under low-temperature stress in the banana. *Plants (Basel)*, 12(16), 2926. doi: 10.3390/plants12162926

García, A.V. et al., 2021. Genomic affinity in hybrids of B-genome *Arachis* species: new genetic resources toward peanut improvement. *Crop Breeding and Applied Biotechnology*, 21(3), e38292139. doi: 10.1590/1984-70332021v21n3a48

Gasteiger, E. et al., 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Research*, 31(13), pp.3784-3788. doi: 10.1093/nar/gkg563

Gasteiger, E. et al., 2005. Protein identification and analysis tools on the ExPASy server. In *The proteomics protocols handbook*. Humana Press, pp.571-607. doi: 10.1385/1-59259-890-0:571

Group, M., 2002. Mitogen-activated protein kinase cascades in plants: a new nomenclature. *Trends in Plant Science*, 7(7), pp.301-308. doi: 10.1016/s1360-1385(02)02302-6

Hall, T.A., 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, 41, pp.95-98. doi: 10.14601/Phytopathol_Mediterr-14998u1.29

Hill, G.M., 2002. Peanut by-products fed to cattle. *The Veterinary clinics of North America: Food animal practice*, 18(2), pp.295-315. doi: 10.1016/s0749-0720(02)00019-1

Hu, B. et al., 2015. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8), pp.1296-1297. doi: 10.1093/bioinformatics/btu817

Jagodzik, P. et al., 2018. Mitogen-Activated Protein Kinase Cascades in Plant Hormone Signaling. *Frontiers in Plant Science*, 9, 1387. doi: 10.3389/fpls.2018.01387

Jonwal, S. et al., 2023. Regulation of photosynthesis by mitogen-activated protein kinase in rice: antagonistic adjustment by OsMPK3 and OsMPK6. *Physiology and Molecular Biology of Plants*, 29(9), pp.1247-1259. doi: 10.1007/s12298-023-01383-9

Kong, F. et al., 2012. Genome-wide analysis of the mitogen-activated protein kinase gene family in *Solanum lycopersicum*. *Gene*, 499(1), pp.108-120. doi: 10.1016/j.gene.2012.01.048

Kosev, V.I. & Vasileva, V.M., 2019. Morphological characterization of Grass pea (*Lathyrus sativus* L.) Varieties. *Journal of Agricultural Sciences - Sri Lanka*, 14(2), pp.67-76. doi: 10.4038/jas.v14i2.8509

La, H.V. et al., 2022a. SWEET gene family in sugar beet (*Beta vulgaris*): Genome-wide survey, phylogeny and expression analysis. *Pakistan journal of biological sciences*, 25(5), pp.387-395. doi: 10.3923/pjbs.2022.387.395

La, H.V. et al., 2022b. Insights into the gene and protein structures of the CaSWEET family members in chickpea (*Cicer arietinum*), and their gene expression patterns in different organs under various stress and abscisic acid treatments. *Gene*, 819, 146210. doi: 10.1016/j.gene.2022.146210

Le, M.T. et al., 2022. Genome-wide identification and analysis of genes encoding putative heat shock protein 70 in papaya (*Carica papaya*). *Pakistan journal of biological sciences*, 25(6), pp.468-475. doi:10.3923/pjbs.2022.468.475

Li, M. et al., 2022. Genome-wide identification and expression of MAPK gene family in cultivated strawberry and their involvement in fruit developing and ripening. *International Journal of Molecular Sciences*, 23(9), 5201. doi: 10.3390/ijms23095201

Li, X. et al., 2014. Identification of rapidly induced genes in the response of peanut (*Arachis hypogaea*) to water deficit and abscisic acid. *BMC Biotechnology*, 14, 58. doi:10.1186/1472-6750-14-58

Lin, L. et al., 2021. Plant mitogen-activated protein kinase cascades in environmental stresses. *International Journal of Molecular Sciences*, 22(4), 1543. doi: 10.3390/ijms22041543

Liu, Y. et al., 2013. Genome-wide analysis of mitogen-activated protein kinase gene family in maize. *Plant Molecular Biology Reporter*, 31(6), pp.1446-1460. doi: 10.1007/s11105-013-0623-y

Majeed, Y. et al., 2022. Functional analysis of mitogen-activated protein kinases (MAPKs) in potato under biotic and abiotic stress. *Molecular Breeding*, 42(6), 31. doi: 10.1007/s11032-022-01302-y

Neupane, A. et al., 2013. Evolutionary history of mitogen-activated protein kinase (MAPK) genes in *Lotus*, *Medicago*, and *Phaseolus*. *Plant Signaling & Behavior*, 8(11), e27189. doi: 10.4161/psb.27189

Niekerk, L.A. et al., 2024. Heavy metal stress and mitogen activated kinase transcription factors in plants: Exploring heavy metal-ROS influences on plant signalling pathways. *Plant, Cell & Environment*, 47(8), pp.2793-2810. doi: 10.1111/pce.14926

Opdenakker, K. et al., 2012. Mitogen-Activated Protein (MAP) kinases in plant metal stress: regulation and responses in comparison to other biotic and abiotic stresses. *International Journal of Molecular Sciences*, 13(6), pp.7828-7853. doi: 10.3390/ijms13067828

Reyna, N.S. & Yang, Y., 2006. Molecular analysis of the rice MAP kinase gene family in relation to *Magnaporthe grisea* infection. *Molecular Plant-Microbe Interactions*, 19(5), pp.530-540. doi: 10.1094/mpmi-19-0530

Singh, A. et al., 2018. Genome-wide identification of the MAPK gene family in chickpea and expression analysis during development and stress response. *Plant Gene*, 13, pp.25-35. doi:10.1016/j.plgene.2017.12.001

Song, A. et al., 2018. Comprehensive analysis of mitogen-activated protein kinase cascades in *Chrysanthemum*. *PeerJ*, 6, e5037. doi: 10.7717/peerj.5037

Taj, G. et al., 2010. MAPK machinery in plants: recognition and response to different stresses through multiple signal transduction pathways. *Plant Signaling & Behavior*, 5(11), pp.1370-1378. doi: 10.4161/psb.5.11.13020

Tamura, K. et al., 2021. MEGA11: Molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution*, 38(7), pp.3022-3027. doi: 10.1093/molbev/msab120

Thompson, J. et al., 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*, 25, pp.4876 - 4882.

Thompson, J.D. et al., 2002. Multiple sequence alignment using ClustalW and ClustalX. *Current Protocols in Bioinformatics*, Chapter 2, Unit 2.3. doi: 10.1002/0471250953.bi0203s00

Toomer, O.T., 2018. Nutritional chemistry of the peanut (*Arachis hypogaea*). *Critical Reviews in Food Science and Nutrition*, 58(17), pp.3042-3053. doi: 10.1080/10408398.2017.1339015

Wadood, S.A. et al., 2022. Geographical origin classification of peanuts and processed fractions using stable isotopes. *Food Chemistry: X*, 16, 100456. doi: 10.1016/j.fochx.2022.100456

Wang, G. et al., 2014. Genome-wide identification and analysis of mitogen activated protein kinase kinase kinase gene family in grapevine (*Vitis vinifera*). *BMC Plant Biology*, 14, 219. doi: 10.1186/s12870-014-0219-1

Wang, J. et al., 2015. Genome-wide identification of MAPK, MAPKK, and MAPKKK gene families and transcriptional profiling analysis during development and stress response in cucumber. *BMC Genomics*, 16(1), 386. doi: 10.1186/s12864-015-1621-2

Wang, T. et al., 2022. Genome-wide identification and expression analysis of MAPK gene family in lettuce (*Lactuca sativa* L.) and functional analysis of LsMAPK4 in high- temperature-induced bolting. *International Journal of Molecular Sciences*, 23(19). doi: 10.3390/ijms231911129

Zhang, M. & Zhang, S., 2022. Mitogen-activated protein kinase cascades in plant signaling. *Journal of Integrative Plant Biology* , 64(2), pp.301-341. doi: 10.1111/jipb.13215

Zheng, Y.M. et al., 2019. Nitrogen fixation characteristics of root nodules in different peanut varieties and their relationship with yield. *Ying yong sheng tai xue bao = The journal of applied ecology*, 30(3), pp.961-968. doi: 10.13287/j.1001-9332.201903.019

Zhuang, W. et al., 2019. The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nature Genetics*, 51(5), pp.865-876. doi: 10.1038/s41588-019-0402-2