Journal of Plant Development
Vol. **32**, Dec **2025**: 143-156
doi: 10.47743/jpd.2025.32.1.974
Research Article

ISSN 2065-3158 print / e-ISSN 2066-9917
Available online: www.plant-journal.uaic.ro

INVESTIGATING THE INTERACTOME OF *ARABIDOPSIS THALIANA* MITOGEN-ACTIVATED PROTEIN KINASES: AN INTEGRATIVE APPROACH USING MULTIPLE SOURCES OF EVIDENCE

Sameer KHAN[1], Tim XING[1*]

[1] Department of Biology and Institute of Biochemistry, Carleton University, Ottawa ON – Canada.
[*] Corresponding author. E-mail: tim.xing@carleton.ca, ORCID: 0000-0002-8325-0835

Abstract: *Arabidopsis thaliana* mitogen-activated protein kinase (MPK) signaling network plays a role in various cellular processes. This study integrated protein-protein interaction, genetic interaction, and co-expression data from the STRING, BioGRID, and ATTED-II databases to provide a comprehensive dataset of interactions within the network. The key MPK network components from this set were identified and subjected to functional enrichment analysis, which revealed their involvement in diverse biological processes and pathways. This integrative approach, combining multiple sources of evidence, provides a comprehensive approach for understanding *Arabidopsis thaliana* MPK signaling network. The findings demonstrate the complex regulatory mechanisms that play a role in plant stress responses and development.

Keywords: gene co-expression, genetic interaction, multi-omics, protein-protein interaction, signaling network.

Introduction

Mitogen-activated protein kinase (MAPK) cascades are highly conserved signal transduction pathways that play crucial roles in regulating plant growth, development, and stress responses [MOHANTA & al. 2015; POPESCU & al. 2008; XING & FOROUD, 2021]. These signaling pathways consist of three main tiers of protein kinases: MPK kinase kinases (MAPKKKs), MPK kinases (MKKs), and MPKs, which are activated through sequential phosphorylation events in response to various extracellular and intracellular signals [LEE & al. 2008; POPESCU & al. 2008]. Traditional genetic and biochemical methods have identified MAPKKK/MKK/MPK signaling modules with overlapping roles in controlling cell division, development, hormone signaling and synthesis, and response to abiotic stress and pathogens [POPESCU & al. 2008].

With multiple members at each of the three levels of this kinase cascade, not all combinations occur. First, genetic evidence does not support the mix-and-match formation of MAPK cascades in plants; second, different kinases are assembled into distinct modules by scaffold proteins; third, other mechanisms function for the specificity of each modules or signaling pathway (e.g. spatiotemporal separation of pathways) [XING & FOROUD, 2021]. With the development of various omics analysis tools, the integration of multi-omics data has become an important component of systems biology, allowing for a comprehensive understanding of the molecular mechanisms underlying plant biology [KATAM & al. 2022; XING & FOROUD, 2021]. This integrated approach provides insights into complex biological

Received: 11 August 2025 / ***Revised***: 4 December 2025 / ***Accepted***: 5 December 2025

processes, such as plant responses to environmental stimuli, interactions with other organisms, and the identification of potential traits for crop improvement [MEENA & al. 2017]. The integration of omics datasets has led to the discovery of different plant stress response mechanisms and has improved our understanding of plant biological processes [CRAMER & al. 2011; ZHOU & al. 2022].

To address the current research gaps and advance our understanding of MPK signaling in plants, in this study, we constructed a dataset by integrating various types of omics data, including protein-protein interactions (PPI), genetic interactions, and gene co-expression, to explore the MPK signaling network in *Arabidopsis thaliana* and to identify novel signaling pathways that may play important roles in plant stress responses.

Material and methods

Retrieval of protein-protein interaction data

The STRING Database

Protein-protein interaction data were obtained from the STRING database (version 12.0) using its application programming interface (API). STRING is an extensive resource that integrates both known and predicted protein-protein interactions, including physical interactions and functional associations. The database collects and scores evidence from various sources, such as automated text mining, experiments, computational predictions, and systematic transfers of interaction evidence across organisms. STRING assesses and transfers this information to less-well-studied organisms using hierarchical orthology [SZKLARCZYK & al. 2011, 2019, 2023].

Data collection and processing

For each of the 20 MPKs of *Arabidopsis thaliana*, we conducted a query in STRING, specifying the species using the NCBI Taxonomic ID 3702. The network type parameter was set to the 'Functional Network', which represents the 'Full STRING Network' and consists of both physical and functional associations. The minimum required score was set at a medium confidence score of 0.4. The maximum number of interactions to retrieve for each MPK was set at 500 to obtain a wide range of interactions. Following the retrieval of individual files, the data were consolidated into a single comprehensive file to simplify integration with subsequent genetic interaction and gene co-expression data. This consolidation process involved merging individual files and ensuring that the data were properly formatted and compatible for further analysis.

Retrieval of genetic interaction data

The BioGRID database

Genetic interactions for the MPKs were obtained from BioGRID (version 4.4.232) using the API. BioGRID is a database that curates protein and genetic interactions in various species, including *Arabidopsis thaliana*, based on primary experimental evidence from scientific literature. This ensures the high quality and reliability of the interactions in the database. BioGRID consists of both low-throughput studies and high-throughput datasets [OUGHTRED & al. 2018, 2020; STARK & al. 2006], and allows for a comprehensive overview of genetic interactions.

Data collection and processing

Each MPK was individually queried using BioGRID, with the 'NCBI Taxonomic ID' set to 3702 for *Arabidopsis thaliana*. The obtained data were then filtered to include only genetic interactions, as our focus was on understanding the genetic relationships between MPKs and other genes. The individual files were subsequently combined into a comprehensive file, retaining certain columns and eliminating others based on their relevance to the analysis. This process involved carefully selecting columns that contained essential information for our study and removing any irrelevant data. By doing so, we streamlined the dataset and prepared it for integration with the protein-protein interaction and gene co-expression data.

Retrieval of gene co-expression data

The ATTED-II database

To obtain gene co-expression data for *Arabidopsis thaliana* MPKs, we accessed the ATTED-II database (version 11.1) using the API. ATTED-II is a comprehensive resource that integrates RNA sequencing and microarray data from *Arabidopsis thaliana* and other plant species to predict gene functions and interactions. This resource identifies genes with similar expression profiles, suggesting functional relationships or joint roles in cellular processes [OBAYASHI & al. 2018, 2022].

The ATTED-II database utilizes principal component analysis (PCA) and ensemble calculations for sample balancing, which improves the quality of the co-expression data. Each gene pair in the database is assigned a metric called the 'Logit Score', which serves as an indicator of the strength of their co-expression relationship [OBAYASHI & al. 2018, 2022]. A higher Logit Score suggests a stronger co-expression relationship between the gene pair, implying a closer functional association. Conversely, a lower Logit Score indicates a weaker co-expression relationship, suggesting a less significant functional association.

Data collection and processing

The query parameters were set to retrieve co-expression relationships based on the "Ath-u.c3-0" platform. Additionally, we limited the number of co-expressed genes retrieved for each MPK to 200. After retrieving the co-expression data from ATTED-II, we processed the data to prepare it for integration with the protein-protein interaction and genetic interaction data. This processing step involved formatting the data, ensuring consistency in gene identifiers, and removing duplicate or irrelevant entries. By cleaning and organizing the co-expression data, we ensured its compatibility with other datasets and facilitated the subsequent integration and analysis steps.

Integration and ranking of data

Integration of data

To gain a comprehensive understanding of the potential interactors of MPKs in *Arabidopsis thaliana*, we integrated PPI data from the STRING database, genetic interaction data from the BioGRID database, and gene co-expression data from the ATTED-II database to construct a comprehensive dataset. Each of these data sources provided a specific type of evidence supporting the potential interactions between MPKs and predicted interactors. The integration process involved combining the three datasets based on common gene identifiers. By merging these datasets, we created a unified resource that captures the diverse types of interactions and relationships between MPKs and other genes in *Arabidopsis thaliana*. This

integrated dataset served as the foundation for our subsequent analysis and prioritization of potential MPK interactors.

Ranking and prioritization strategy

To prioritize the relevant MPK interactions in the dataset, we used a ranking strategy that considered the presence of multiple sources of evidence supporting each interaction. This method allowed us to assign a higher degree of confidence to interactions supported by multiple sources. The ranking and prioritization strategies implemented are as follows:

Rank 1: Interactions supported by all three sources of evidence (PPIs from STRING, genetic interactions from BioGRID, and gene co-expression from ATTED-II) were ranked highest. These interactions were considered the most reliable, as they were consistently observed across different sources of evidence.

Rank 2: Interactions supported by two of the three sources of evidence were ranked lower than those supported by all three; however, they were still higher than the interactions supported by only one source of evidence.

Rank 3: Interactions supported by only a single source of evidence were ranked the lowest among all interactions. Although these interactions were considered potentially relevant, they were assigned a lower confidence level than those supported by multiple sources of evidence.

Selection of key MPK network components

After ranking the interactions, we further refined the dataset by selecting the interaction pairs with at least two sources of evidence. This criterion ensured that the selected interactions had a higher level of confidence and were supported by multiple experimental or computational approaches.

To focus on the most critical components of the MPK network, we further narrowed down the dataset to include only key interactions. This selection process involved considering factors such as the biological relevance of the interacting partners, their known functions, and their potential roles in MPK signaling pathways. By concentrating on these key components, we aimed to identify the most essential and influential interactions within the MPK network.

Following the selection of these key interactions, a literature search was conducted to determine whether they were already known or predicted. The results of this literature search were categorized into two groups: known interactions supported by experimental evidence and predicted interactions that had not been experimentally validated. This categorization provided valuable insights into the current state of knowledge regarding MPK interactions and highlighted potential novel interactions that warrant further investigation.

Fundamentals of overrepresentation analysis

For the functional genomics analysis of the selected MPK interactions, we used an overrepresentation-based enrichment approach. Overrepresentation Analysis (ORA) is a statistical method used to determine if a predefined set of genes is represented more than expected by chance within a specific category of genes, such as those belonging to a particular Gene Ontology (GO) term [ASHBURNER & al. 2000]. ORA typically involves the use of a hypergeometric test or similar statistical methods to calculate the probability that the observed overlap between the query gene set and GO category occurs by random chance [ASHBURNER & al. 2000].

Functional and pathway enrichment analysis

g:Profiler (g:GOSt) functional enrichment web server

The functional and pathway enrichment analysis for the predicted MPK interaction genes was performed using g:Profiler (version e111_eg58_p18_30541362), a bioinformatics tool specifically designed for this purpose. g:Profiler allows for the identification of biological pathways that are significantly represented in a given list of genes or proteins [KOLBERG & al. 2023].

Using g:Profiler involves several steps, starting with the input of a gene list derived from either computational or experimental data. The algorithm then determines statistically enriched pathways by comparing the input list to known biological pathways and assessing the overrepresentation of pathways in the gene list more than would be expected by chance. The results can be visualized and interpreted to gain insights into the functional relationships underlying the query genes [KOLBERG & al. 2023].

Statistical settings and data sources

The statistical domain scope was set to all annotated genes, and the significance method selected was g:SCS with a threshold of 0.05, which is the default method for multiple testing correction in g:Profiler. This helps with minimizing false-positive findings and ensuring reliable results [KOLBERG & al. 2023]. The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases were selected as the data sources. For the GO data, the GO Molecular Function (GO:MF), GO Biological Process (GO:BP), and GO Cellular Component (GO:CC) categories were selected. Additionally, electronic GO annotations were included in the enrichment analysis.

Visualization and analysis

To visually represent the distribution of enriched terms for each category, a Manhattan plot was generated using g:Profiler. The Manhattan plot was used to provide a comprehensive overview of the enrichment results, with each point representing a specific term and its corresponding adjusted p-value. The x-axis represents the functional categories (GO, KEGG, and WikiPathways (WP)), while the y-axis represents the negative logarithm of the adjusted p-value. This plot was used to assess the overall enrichment results and identify the most significant functional categories and pathways related to the MPK interactome.

Considering that all input genes were components of the MPK network, it was anticipated that several widely recognized broad enrichments would emerge. However, to gain novel insights into the potential roles and mechanisms of the MPK interactome, it was essential to identify unexpectedly significantly enriched terms, as these unanticipated enrichments might suggest previously uncharacterized functions or regulatory pathways.

To better visualize and interpret these unexpected enrichments, a binary enrichment heatmap was generated using the ggplot2 package in R [WICKHAM, 2016]. The heatmap displays the presence or absence of significantly enriched terms across different functional categories (GO terms, KEGG pathways, and WikiPathways) for each input gene. The binary representation allows for a clear and concise visualization of the enrichment patterns, facilitating the identification of shared and unique functional associations among the MPK-related genes.

Results

Integrated dataset of MPK interactions

The integration of protein-protein interaction data from STRING, genetic interaction data from BioGRID, and gene co-expression data from ATTED-II resulted in a comprehensive dataset containing potential interactors of *Arabidopsis thaliana* MPKs. The final dataset contains the Number of Sources (num_sources), Source Databases (source_db), STRING Score (string_score), ATTED Logit Score (atted_ls), and BioGRID Genetic Evidence Type (biogrid_evidence) for each interaction. For interactions where a specific field was not applicable, it was marked as "N/A". The dataset can be found in the supplementary materials [https://doi.org/10.5281/zenodo.11113250].

Selected MPK network components

The criteria for selecting the MPK network interactions for further analysis were based on the requirement of having a minimum of two sources of evidence and interactors being upstream of MPKs. These selected interactions include the following: MKK2-MPK4, MKK1-MPK1, MKK1-MPK2, MKK2-MPK2, MKK4-MPK3, MKK9-MPK3, MEKK1-MPK3, MKK1-MPK4, MKK2-MPK6, MKK5-MPK6, YDA-MPK9, MKK1-MPK11, MKK6-MPK13, and MEKK1-MPK17. Table 1 presents the selected interactions, displaying the STRING score, ATTED logit score, and BioGRID genetic evidence type for each interaction.

Table 1. Selected subset of upstream MPK interactions with multiple sources of evidence

MPK	Interactor	Number of Sources	Source	STRING Score	ATTED Logit Score	BioGRID Evidence Type
MPK4	MKK2	3	STRING, BioGRID, ATTED	0.993	4.9821	synthetic rescue
MPK1	MKK1	2	STRING, ATTED	0.693	3.344	n/a
MPK2	MKK1	2	STRING, ATTED	0.700	3.5388	n/a
MPK2	MKK2	2	STRING, ATTED	0.659	5.1195	n/a
MPK3	MKK4	2	STRING, ATTED	0.990	4.6728	n/a
MPK3	MKK9	2	STRING, ATTED	0.970	4.4093	n/a
MPK3	MEKK1	2	STRING, ATTED	0.705	3.6991	n/a
MPK4	MKK1	2	STRING, ATTED	0.996	3.8939	n/a
MPK6	MKK2	2	STRING, BioGRID	0.992	N/A	synthetic rescue
MPK6	MKK5	2	STRING, ATTED	0.991	3.8137	n/a
MPK9	YDA	2	STRING, ATTED	0.406	2.9775	n/a
MPK11	MKK1	2	STRING, ATTED	0.88	4.7415	n/a
MPK13	MKK6	2	STRING, ATTED	0.942	4.9706	n/a
MPK17	MEKK1	2	STRING, ATTED	0.459	3.0577	n/a

Experimental evidence and computational predictions

A literature search was conducted to categorize the selected interactions as either known interactions with supporting experimental evidence or computationally predicted interactions lacking experimental validation. The interactions that are known and established through experimental evidence are as follows: MKK2-MPK4, MKK2-MPK2, MKK4-MPK3, MKK9-MPK3, MKK1-MPK4, MKK2-MPK6, MKK5-MPK6, and MKK1-MPK11. These known interactions are presented in Table 2, which includes their associated functions according to the literature.

Table 2. Experimentally validated upstream interactions within the selected subset

MPK	Interactor	Functions	References
MPK4	MKK2	cold and salt stress signaling, represses cell death and immune response	KONG & al. (2012) and TEIGE & al. (2004)
MPK2	MKK2	positive regulation of red light-induced stomatal opening	LI & al. (2023)
MPK3	MKK4	Agrobacterium-triggered immunity, Agrobacterium-mediated transformation	LIU & al. (2021)
MPK3	MKK9	enhances phosphate acquisition, ethylene and camalexin biosynthesis, enhances salt stress sensitivity	LEI & al. (2014) XU & al. (2008)
MPK4	MKK1	negatively regulates immunity	KONG & al. (2012)
MPK6	MKK2	cold and salt stress signaling	TEIGE & al. (2004)
MPK6	MKK5	ABA regulation of primary root growth, stomatal response	LI & al. (2017)
MPK11	MKK1	n/a	LEE & al. (2008)

The predicted interactions, which lack experimental evidence but were identified through computational methods, are as follows: MKK1-MPK1, MKK1-MPK2, MEKK1-MPK3, YDA-MPK9, MKK6-MPK13, and MEKK1-MPK17. These computationally predicted interactions are presented in Table 3, including their associated scores from STRING and ATTED-II.

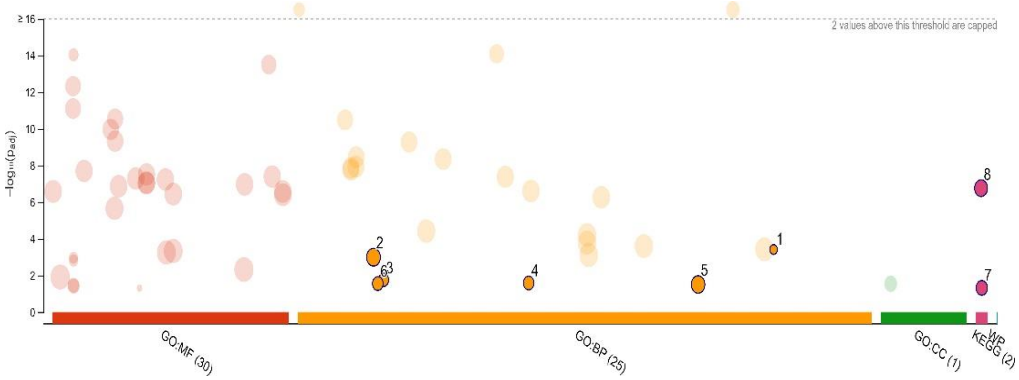
Table 3. Computationally predicted interactions within the selected subset

MPK	Interactor	Sources	STRING Score	ATTED Logit Score
MPK1	MKK1	STRING, ATTED	0.693	3.344
MPK2	MKK1	STRING, ATTED	0.700	3.5388
MPK3	MEKK1	STRING, ATTED	0.705	3.6991
MPK9	YDA	STRING, ATTED	0.406	2.9775
MPK13	MKK6	STRING, ATTED	0.942	4.9706
MPK17	MEKK1	STRING, ATTED	0.459	3.0577

Pathway enrichment analysis

The functional and pathway enrichment analysis of the predicted MPK interaction genes revealed a number of significantly enriched categories, consisting of GO terms and KEGG pathways. The Manhattan plot displayed in Figure 1 illustrates the distribution of the enrichment results.

Figure 1. Manhattan plot illustrating the functional and pathway enrichment results



Among the significantly enriched terms identified, many were anticipated based on prior knowledge. To focus the analysis on the most relevant and intriguing findings, terms that are particularly noteworthy were carefully selected. These selected terms of interest, along with their associated genes and adjusted p-values, are presented in Table 4. The dataset is saved as a .txt file and can be accessed at https://doi.org/10.5281/zenodo.11113422.

Table 4. Selected enriched terms, adjusted p-values, and the associated genes

Term Description	Adjusted P-Value	Genes
Response to L-glutamate	3.85×10^{-4}	MPK3, MEKK1
Response to wounding	1.01×10^{-3}	MPK2, MPK3, MKK1, MEKK1
Inflorescence development	1.76×10^{-2}	MPK3, YDA
Response to amino acid	2.59×10^{-2}	MPK3, MEKK1
Pollen-pistil interaction	2.78×10^{-2}	MPK3, MKK1
Post-embryonic plant morphogenesis	3.12×10^{-2}	MPK13, YDA, MKK6
Efferocytosis	4.79×10^{-2}	MPK13, MKK6

Discussion

Significance of the integrated MPK interaction dataset

The integration of multiple sources of evidence, such as protein-protein interactions, genetic interactions, and gene co-expression data, is crucial for exploring potential interactions within the *Arabidopsis thaliana* MPK signaling network. This approach enhances the reliability

and accuracy of predictions by providing a comprehensive understanding of the complex nature of cellular processes [WU & al. 2010]. By integrating diverse biological and computational sources of evidence, we can mitigate the limitations of individual data sources, such as high noise levels in high-throughput PPI data and inherent variability in gene expression profiles [SRISWASDI & JENSEN, 2012; TU & al. 2006].

The integrated dataset offers several advantages over single-data-source approaches. For instance, gene expression data can refine the topology of PPI networks by removing less relevant interactions, thereby simplifying the interactome for improved biological coherence [SRISWASDI & JENSEN, 2012]. Additionally, the condition-specific nature of PPIs suggests that integrating data on transcriptional regulation and gene expression can provide insights into the dynamic behavior of protein complexes under different environmental conditions [LUO & al. 2010]. The integration of multiple sources of evidence compensates for the weaknesses of individual data types and provides a more robust framework for the prediction and analysis of PPIs and genetic interactions, leading to improved prediction coverage and accuracy [LI & al. 2012; WU & al. 2010]. There are several potential applications of this dataset. By providing a comprehensive view of the *Arabidopsis thaliana* MPK signaling network, this dataset can help address current research gaps, such as the characterization of uncharacterized MPKs and understanding of MAPKKK gene functions in plant stress responses. The integrated dataset can serve as a valuable resource for researchers investigating the roles of MPKs in various cellular processes and can guide future experimental studies aimed at validating predicted interactions and uncovering novel functional relationships within the MPK signaling network.

Characteristics of the selected MPK network components

The selection of key MPK network components from the integrated dataset was based on the presence of multiple sources of evidence supporting each interaction. Interactions supported by all three sources of evidence (PPIs from STRING, genetic interactions from BioGRID, and gene co-expression from ATTED-II) were ranked highest, followed by those supported by two sources, and finally those supported by a single source. This ranking strategy allowed us to prioritize the interactions based on the strength and consistency of the evidence supporting them, focusing on the most promising and reliable interactions for further investigation and validation.

The confidence scores associated with PPIs in the STRING database served as quantitative indicators of the strength of evidence for each predicted interaction. Higher scores suggest a greater likelihood of true biological significance, allowing for the prioritization of interactions for further investigation [SZKLARCZYK & al. 2011, 2019, 2023]. Similarly, the logit score used in the ATTED-II database to measure gene co-expression provides a reliable assessment of the strength of co-expression between genes, with higher scores indicating stronger co-expression relationships [OBAYASHI & al. 2018, 2022]. The combined co-expression platform, which integrates both microarray and RNA-Seq data, has been shown to have a higher function score than individual platforms, indicating better predictive performance [OBAYASHI & al. 2018, 2022].

Genetic interaction data from the BioGRID database was limited for *Arabidopsis thaliana* MPKs, with only two interactions (MKK2-MPK4 and MKK2-MPK6) identified in a previous study [TEIGE & al. 2004]. These interactions were categorized as “Synthetic Rescue”, referring to the rescue of lethality or growth defects in a strain mutated/deleted for one gene by the mutation/deletion of another gene [OUGHTRED & al. 2018, 2020; STARK, 2006]. The limited availability of genetic interaction data for MPKs in *Arabidopsis thaliana* highlights the

need for further experimental studies to uncover additional genetic relationships within the MPK signaling network.

The distribution of evidence types among the selected MPK network components has important implications for the reliability and biological relevance of the predicted interactions. Interactions supported by multiple sources of evidence are more likely to represent true functional relationships, as they have been consistently observed across different experimental and computational approaches. By focusing on these high confidence interactions, we can identify the most essential and influential components of the MPK signaling network and guide future research efforts towards understanding their roles in plant stress responses and development.

Experimentally validated interactions and biological roles

In this exploratory study, we identified several experimentally validated interactions among selected interactions in *Arabidopsis thaliana*. These interactions are supported by various experimental approaches that provide valuable insights into their functional roles and biological significance.

(1) MKK2-mediated interactions and their role in cold and salt stress signaling, immunity, and red light-induced stomatal opening

The interactions mediated by MKK2 are important for the regulation of stress signaling, immunity, and stomatal opening. Specifically, MKK2 activates MPK4, a process essential for plants to respond to cold and salt stress, as well as the negative regulation of plant immunity, which helps suppress unnecessary cell death and immune responses under normal conditions [KONG & al. 2012; TEIGE & al. 2004].

The MKK2-MPK2 interaction plays an important role in red light-induced stomatal opening, with MKK2 phosphorylating MPK2 in guard cells to trigger stomatal opening [LI & al. 2023]. Additionally, the interaction of MKK2 with MPK6 contributes to the adaptation of plants to cold and salt stress [TEIGE & al. 2004], highlighting the diverse roles of MKK2 in various physiological processes.

(2) MKK4 and MKK9-mediated interactions and their role in agrobacterium-triggered immunity, phosphate acquisition, and stress responses

The MKK4-MPK3 interaction is important for regulating plant immunity and transforming responses to Agrobacterium infection, inducing defense-responsive gene expression [LIU & al. 2021]. The MKK9-MPK3 interaction enhances phosphate acquisition by regulating the transcription of phosphate-responsive genes and activates MPK3, influencing ethylene and camalexin biosynthesis as well as salt stress responses [LEI & al. 2014; XU & al. 2008].

(3) MKK1 and MKK5-mediated interactions and their role in suppressing immunity and ABA-regulated root growth and stomatal responses

The MKK1-MPK4 interaction is a key element of the MPK signaling cascade that suppresses immune responses to prevent autoimmunity [KONG & al. 2012]. Additionally, the MKK5-MPK6 interaction, as part of the AIK1-MKK5-MPK6 cascade, plays a critical role in ABA regulation of root growth and stomatal responses, and is involved in regulating stomatal development, ethylene signaling, nitric oxide production, and hydrogen peroxide responses during plant growth and development [LI & al. 2017].

(4) Limited biological significance of the MKK1-MPK11 interaction

Although the physical interaction between MKK1 and MPK11 has been determined using yeast two-hybrid screening, the biological function of this interaction appears to be limited. *In vitro* phosphorylation assays showed very weak activity of MKK1 with MPK11 as a substrate

[LEE & al. 2008], suggesting that the functional significance of this interaction may be less prominent than that of other validated interactions.

Indications of biological roles from computational prediction

(1) MPK signaling in response to nutrient sensing

Our computational predictions have identified significant interactions within the MPK signaling pathways, particularly involving MPK3 and MEKK1 in the response to L-glutamate and amino acids. These results suggest a potential link between MPK signaling and the sensing or metabolism of these molecules. L-glutamate serves as a metabolic signaling molecule in plants, potentially activating the MEKK1-MPK3 signaling to influence nutrient sensing and signaling mechanisms compounds [FORDE & LEA, 2007; KAN & al. 2017; MCCOY & al. 2020].

(2) MPK pathways in wound response

The involvement of MPK2, MPK3, MKK1, and MEKK1 in response to wounding highlights a robust activation of MPK cascades following physical damage. This finding aligns with previous research, suggesting that MPK signaling is crucial for initiating hormonal signaling and transcriptional reprogramming during defense responses and tissue repair. [SÖZEN & al. 2020].

(3) Developmental roles of MPK cascades

The MPK signaling network appears to play an important role in various developmental processes. The association of MPK3 with YDA, and the involvement of the YDA-MPK9 module, might be related to their regulatory roles in cellular patterning and differentiation within floral tissues [ZHANG & ZHANG, 2022]. Our analysis also indicates a role for MPK3 and MKK1 in pollen-pistil interactions, which are important for successful reproduction [LI & al. 2013]. This interaction may influence pollen tube growth and guidance, highlighting the integral role of MPK signaling in facilitating communication between pollen and pistil tissues.

(4) Novel roles in plant cell death

While the direct link between plant MAPKs and efferocytosis (the process of engulfing and clearing apoptotic cells) isn't extensively studied, MAPKs are known to be involved in related processes like programmed cell death and autophagy [ZHANG & ZHANG, 2022], both of which can be linked to efferocytosis. Interestingly, in this work, the involvement of MPK13 and MKK6 in “efferocytosis” (Table 4) opens new avenues for research into the clearance of apoptotic cells, a relatively unexplored area in plant biology. This finding suggests a potential regulatory role for these kinases in managing plant cell death and survival.

Conclusions

The mitogen-activated protein kinase signaling network plays a pivotal role in regulating plant stress responses and development. Our work has employed a systematic and integrative approach to exploring the complex interactions within the *Arabidopsis thaliana* MPK network. Through the integration of protein-protein interactions, genetic interactions, and gene co-expression data, a comprehensive dataset of potential MPK interactors was constructed and analyzed. This integrative strategy enhanced the reliability of predictions by leveraging multiple lines of evidence, compensating for the limitations of individual data sources.

The functional genomics analysis of the predicted MPK interactors revealed novel potential roles for MPK cascades in nutrient sensing, wound response, developmental processes, reproductive biology, and plant cell death regulation. These findings contribute to our

understanding of the intricate regulatory mechanisms underlying the MPK signaling network, and establish a foundation for future experimental validation, mutagenesis studies, and the development of strategies to enhance plant stress tolerance.

Acknowledgements

This work was supported by a research grant to Tim XING from Bayer Crop Science Inc. and Carleton University, Canada.

References

ASHBURNER M., BALL C. A, BLAKE J. A., BOTSTEIN D., BUTLER H., CHERRY J. M., DAVIS A. P., DOLINSKI K., DWIGHT S. S., EPPIG J. T., HARRIS M. A., HILL D. P., ISSEL-TARVER L., KASARSKIS A., LEWIS S., MATESE J. C., RICHARDSON J. E., RINGWALD M., RUBIN G. M. & SHERLOCK G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*. **25**(1): 25-29. https://doi.org/10.1038/75556

CRAMER G. R., URANO K., DELROT S., PEZZOTTI M. & SHINOZAKI K. 2011. Effects of abiotic stress on plants: a systems biology perspective. *BMC Plant Biology*. **11**: 163. https://doi.org/10.1186/1471-2229-11-163

FORDE B. G. & LEA P. J. 2007. Glutamate in plants: metabolism, regulation, and signaling. *Journal of Experimental Botany*. **58**(9): 2339-2358. https://doi.org/10.1093/jxb/erm121

KAN C. C., CHUNG T. Y., WU H. Y., JUO Y. A. & HSIEH M. H. 2017. Exogenous glutamate rapidly induces the expression of genes involved in metabolism and defense responses in rice roots. *BMC Genomics*. **18**(1): 186. https://doi.org/10.1186/s12864-017-3588-7

KATAM R., LIN C., GRANT K., KATAM C. & CHEN S. 2022. Advances in plant metabolomics and its applications in stress and single-cell biology. *International Journal of Molecular Sciences*. **23**(13): 6985. https://doi.org/10.3390/ijms23136985

KOLBERG L., RAUDVERE U., KUZMIN I., ADLER P., VILO J. & PETERSON H. 2023. g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*. **51**(W1): W207-W212. https://doi.org/10.1093/nar/gkad347

KONG Q., QU N., GAO M., ZHANG Z., DING X., YANG F., LI Y., DONG O. X, CHEN S., LI X. & ZHANG Y. 2012. The MEKK1-MKK1/MKK2-MPK4 kinase cascade negatively regulates immunity mediated by a mitogen-activated protein kinase kinase kinase in *Arabidopsis*. *Plant Cell*. **24**(5): 2225-2236. https://doi.org/10.1105/tpc.112.097253

LEE J. S., HUH K. W., BHARGAVA A. & ELLIS B. E. 2008. Comprehensive analysis of protein-protein interactions between Arabidopsis MAPKs and MAPK kinases helps define potential MAPK signalling modules. *Plant Signaling & Behavior*. **3**(12): 1037-1041. https://doi.org/10.4161/psb.3.12.6848

LEI L., LI Y., WANG Q., XU J., CHEN Y., YANG H. & REN D. 2014. Activation of MKK9-MPK3/MPK6 enhances phosphate acquisition in *Arabidopsis thaliana*. *New Phytologist*. **203**(4): 1146-1160. https://doi.org/10.1111/nph.12872

LI M., WU X., WANG J. & PAN Y. 2012. Towards the identification of protein complexes and functional modules by integrating PPI network and gene expression data. *BMC Bioinformatics*. **13**(1): 109. https://doi.org/10.1186/1471-2105-13-109

LI X. M., SANG Y. L., ZHAO X. Y. & ZHANG X. S. 2013. High-throughput sequencing of small RNAs from pollen and silk and characterization of miRNAs as candidate factors involved in pollen-silk interactions in maize. *PLoS ONE*. **8**: e72852. https://doi.org/10.1371/journal.pone.0072852

LI K., YANG F., ZHANG G., SONG S., LI Y., REN D., MIAO Y. & SONG C. P. 2017. AIK1, a mitogen-activated protein kinase, modulates abscisic acid responses through the MKK5-MPK6 kinase cascade. *Plant Physiology*. **173**(2): 1391-1408. https://doi.org/10.1104/pp.16.01386

LI Y., HUI S., YUAN Y., YE Y., MA X., ZHANG X., ZHANG S., ZHANG C. & CHEN Y. P. 2023. phyB-dependent phosphorylation of mitogen-activated protein kinase cascade MKK2-MPK2 positively regulates red light-induced stomatal opening. *Plant Cell and Environment*. **46**(11): 3323-3336. https://doi.org/10.1111/pce.14675

LIU T., CAO L., CHENG Y., JI J., WEI Y., WANG C. & DUAN K. 2021. MKK4/5-MPK3/6 cascade regulates *Agrobacterium*-mediated transformation by modulating plant immunity in *Arabidopsis*. *Frontiers in Plant Science*. **12**: 731690. https://doi.org/10.3389/fpls.2021.731690

LUO F., LIU J. & LI J. 2010. Discovering conditional co-regulated protein complexes by integrating diverse data sources. *BMC Systems Biology*. **4**(Suppl 2): S4. https://doi.org/10.1186/1752-0509-4-S2-S4

MCCOY R. M., MEYER G. W., RHODES D., MURRAY G. C., SORS T. G. & WIDHALM J. R. 2020. Exploratory study on the foliar incorporation and stability of isotopically labeled amino acids applied to turfgrass. *Agronomy.* **10**(3): 358. https://doi.org/10.3390/agronomy10030358

MEENA K. K., SORTY A. M., BITLA U. M., CHOUDHARY K., GUPTA P., PAREEK A., SINGH D. P., PRABHA R., SAHU P. K., GUPTA V. K., SINGH H. B., KRISHANANI K. K. & MINHAS P. S. 2017. Abiotic stress responses and microbe-mediated mitigation in plants: the omics strategies. *Frontiers in Plant Science*. **8**: 172. https://doi.org/10.3389/fpls.2017.00172

MOHANTA T. K., MOHANTA N., PARIDA P. & BAE H. 2015. Mitogen activated protein kinase (MPK) interacts with auxin influx carrier (OsAux/LAX1) involved in auxin signaling in plant. *Biological Procedures Online*. **17**: 13. https://doi.org/10.1186/s12575-015-0025-7

OBAYASHI T., AOKI Y., TADAKA S., KAGAYA Y. & KINOSHITA K. 2018. ATTED-II in 2018: A plant co-expression database based on investigation of the statistical property of the mutual rank index. *Plant and Cell Physiology*. **59**(2): 440. https://doi.org/10.1093/pcp/pcx209

OBAYASHI T., HIBARA H., KAGAYA Y., AOKI Y. & KINOSHITA K. 2022. ATTED-II v11: a plant gene co-expression database using a sample balancing technique by subagging of principal components. *Plant and Cell Physiology*. **63**(6): 869-881. https://doi.org/10.1093/pcp/pcac041

OUGHTRED R., STARK C., BREITKREUTZ B. J., RUST J., BOUCHER L., CHANG C., KOLAS N., O'DONNELL L., LEUNG G., MCADAM R., ZHANG F., DOLMA S., WILLEMS A., COULOMBE-HUNTINGTON J., CHATR-ARYAMONTRI A., DOLINSKI K. & TYERS M. 2019. The BioGRID interaction database: 2019 update. *Nucleic Acids Research*. **47**(D1): D529-D541. https://doi.org/10.1093/nar/gky1079

OUGHTRED R., RUST J., CHANG C., BREITKREUTZ B. J., STARK C., WILLEMS A., BOUCHER L., LEUNG G., KOLAS N., ZHANG F., DOLMA S., COULOMBE-HUNTINGTON J., CHATR-ARYAMONTRI A., DOLINSKI K. & TYERS M. 2021. The BioGRID database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science*. **30**(1): 187-200. https://doi.org/10.1002/pro.3978

POPESCU S. C., POPESCU G. V., BACHAN S., ZHANG Z., GERSTEIN M., SNYDER M. & DINESH-KUMAR S. P. 2009. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes & Development*. **23**(1): 80-92. https://doi.org/10.1101/gad.1740009

SÖZEN C., SCHENK S. T., BOUDSOCQ M., CHARDIN C., ALMEIDA-TRAPP M., KRAPP A., HIRT H., MITHÖFER A. & COLCOMBET J. 2020. Wounding and insect feeding trigger two independent MAPK pathways with distinct regulation and kinetics. *Plant Cell*. **32**(6): 1988-2003. https://doi.org/10.1105/tpc.19.00917

SRISWASDI S. & JENSEN S. T. 2012. Refining the protein-protein interactome using gene expression data. *arXiv*:1003.2811. https://doi.org/10.48550/arXiv.1003.2811

STARK C., BREITKREUTZ B. J., REGULY T., BOUCHER L., BREITKREUTZ A. & TYERS M. 2006. BioGRID: a general repository for interaction datasets. *Nucleic Acids Research*. **34**(D1): D535-539. https://doi.org/10.1093/nar/gkj109

SZKLARCZYK D., FRANCESCHINI A., KUHN M., SIMONOVIC M., ROTH A., MINGUEZ P., DOERKS T., STARK M., MULLER J., BORK P., JENSEN L. J. & VON MERING C. 2011. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Research.* **39** (D1): D561-568. https://doi.org/10.1093/nar/gkq973

SZKLARCZYK D., GABLE A. L., LYON D., JUNGE A., WYDER S., HUERTA-CEPAS J., SIMONOVIC M., DONCHEVA N. T., MORRIS J. H., BORK P., JENSEN L. J. & MERING C. V. 2019. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*. **47**(D1): D607-D613. https://doi.org/10.1093/nar/gky1131

SZKLARCZYK D., KIRSCH R., KOUTROULI M., NASTOU K., MEHRYARY F., HACHILIF R., GABLE A. L., FANG T., DONCHEVA N. T., PYYSALO S., BORK P., JENSEN L. J. & VON MERING C. 2023. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Research*. **51**(D1): D638-D646. https://doi.org/10.1093/nar/gkac1000

TEIGE M., SCHEIKL E., EULGEM T., DÓCZI R., ICHIMURA K., SHINOZAKI K., DANGL J. L. & HIRT H. 2004. The MKK2 pathway mediates cold and salt stress signaling in *Arabidopsis*. *Molecular Cell*. **15**(1): 141-152. https://doi.org/10.1016/j.molcel.2004.06.023

TU K., YU H. & LI Y. X. 2006. Combining gene expression profiles and protein-protein interaction data to infer gene functions. *Journal of Biotechnology*. **124**(3): 475-485. https://doi.org/10.1016/j.jbiotec.2006.01.024

WICKHAM H. 2016. ggplot2: Elegant graphics for data analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4. https://ggplot2.tidyverse.org

WU M., LI X., CHUA H. N., KWOH C. K. & NG S. K. 2010. Integrating diverse biological and computational sources for reliable protein-protein interactions. *BMC Bioinformatics*. **11**(Suppl 7): S8. https://doi.org/10.1186/1471-2105-11-S7-S8

XING T. & FOROUD N. A. 2021. A network of MAP kinase pathways in pant defense regulation: what we have learned from genetic modification studies? Pp. 393-407. In: PANDEY G. K. (ed.). *Protein kinases and stress signaling in plants: Functional genomic perspective*. John Wiley & Sons Inc.

XU J., LI Y., WANG Y., LIU H., LEI L., YANG H., LIU G. & REN D. 2008. Activation of MAPK kinase 9 induces ethylene and camalexin biosynthesis and enhances sensitivity to salt stress in *Arabidopsis*. *Journal of Biological Chemistry*. **283**(40): 26996-27006. https://doi.org/10.1074/jbc.M801392200

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

ZHOU R., JIANG F., NIU L., SONG X., YU L., YANG Y. & WU Z. 2022. Increase crop resilience to heat stress using omic strategies. *Frontiers in Plant Science*. **13**: 891861. https://doi.org/10.3389/fpls.2022.891861

How to cite this article:

KHAN S. & XING T. 2025. Investigating the interactome of *Arabidopsis thaliana* mitogen-activated protein kinases: an integrative approach using multiple sources of evidence. *J. Plant Develop*. **32**: 143-156. https://doi.org/10.47743/jpd.2025.32.1.974