



OPEN ACCESS

EDITED BY

Sofia R. Pauleta,
New University of Lisbon, Portugal

REVIEWED BY

Divya Prakash,
Southern Illinois University Carbondale,
United States
Sumedha Dahal,
Memorial Sloan Kettering Cancer Center,
United States

*CORRESPONDENCE

Dafeng Liu,
✉ dafeli@sina.cn,
✉ dafeli-dafeli@hotmail.com

RECEIVED 13 May 2025

ACCEPTED 28 June 2025

PUBLISHED 10 July 2025

CITATION

Liu D, Li N, Song D and Lv Z (2025) Structural bioinformatics and gene expression analysis of maturase K from *Lavandula angustifolia* (lavender). *Front. Mol. Biosci.* 12:1628118. doi: 10.3389/fmolb.2025.1628118

COPYRIGHT

© 2025 Liu, Li, Song and Lv. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Structural bioinformatics and gene expression analysis of maturase K from *Lavandula angustifolia* (lavender)

Dafeng Liu^{1,2*}, Na Li¹, Daoqi Song² and Zhenming Lv²

¹Xinjiang Key Laboratory of Lavender Conservation and Utilization, College of Biological Sciences and Technology, Yili Normal University, Yili, China, ²School of Life Sciences, Xiamen University, Xiamen, China

The chloroplast genome of plants contains a single gene encoding the splicing factor Maturase K (MatK). To elucidate the functional role and underlying mechanism of MatK, we investigated it in *Lavandula angustifolia* (lavender). Structural models of MatK1 and MatK2 were predicted using AlphaFold2, and potential active site residues were identified via the GalaxyWEB program. The results of RT-qPCR analysis revealed that the expression of *MatK1* and *MatK2* peaked in leaves at 14:00. For heat treatments, *MatK1* expression in leaves increased with the duration of heat exposure, reaching its highest levels at 40°C for 3 h and 30°C for 6 h, before declining. Similarly, under salt treatment, *MatK1* expression in leaves showed an increasing trend with exposure time, peaking at 300 mM NaCl for 3 h and 200 mM for 12 h, before decreasing. This study provides the first detailed characterization of Maturase K in *L. angustifolia*.

KEYWORDS

lavandula x intermedia (lavandin), maturase K, prediction of structural models, RT-qPCR analysis, heat and salt stress

Introduction

Lavender plants are compact, aromatic shrubs widely cultivated for their essential oils (EOs), which consist of complex blends of mono- and sesquiterpenoid alcohols, esters, oxides, and ketones (Crişan et al., 2023; de Melo Alves Silva et al., 2023). The *Lavandula* genus includes 30 recognized species, with *Lavandula angustifolia*, *Lavandula latifolia*, and *Lavandula x intermedia*—a natural hybrid of *L. latifolia* and *L. angustifolia*—being of significant economic importance (Crişan et al., 2023; de Melo Alves Silva et al., 2023; Landmann et al., 2007; Liu et al., 2025d; Liu et al., 2025c). The highest-quality EOs are derived from the flowering tops of *L. angustifolia*, commonly known as ‘true lavender’, which has been valued for its distinctive fragrance since ancient times. Lavender EOs have diverse applications in cosmetics, hygiene, and alternative medicine (Hedayati et al., 2024; Khan et al., 2024; Li et al., 2024). For example, EOs with elevated camphor concentrations are used in inhalants to treat respiratory conditions like coughs and colds, as well as in liniments and balms for topical pain relief (Malloggi et al., 2021; Batiha et al., 2023; Braunstein and Braunstein, 2023; Liu et al., 2024a). Camphor has also been investigated as a radiosensitizing agent to enhance tumor oxygenation prior to radiotherapy (Bungau et al., 2023; Crişan et al., 2023; de Melo Alves Silva et al., 2023; Dewanjee et al., 2023; Khan et al., 2024).

The production of EOs in plants is closely linked to photosynthesis, a process involving several enzymes, including Maturase K (MatK). Recently, MatK has gained attention as a crucial gene due to its strong phylogenetic signal (Mukhopadhyay and Hausner, 2024). The high rate of amino acid substitution in MatK is attributed to the nearly uniform distribution of substitution rates across the three codon positions, in contrast to most protein-coding genes, where substitution rates are typically biased toward the third codon position (Unnikrishnan et al., 2021; Zhang et al., 2021; Algarni, 2022). In addition to its significance in plant phylogenetics, MatK is the only putative group II intron maturase encoded in the chloroplast genome. MatK enzymes catalyze the nonautocatalytic removal of introns from precursor RNAs. These enzymes typically consist of three domains: a reverse-transcriptase domain, domain X (the proposed maturase functional domain), and a zinc-finger-like domain. While there is a substantial body of literature on MatK in plants (Mukhopadhyay and Hausner, 2024; Tripodi, 2023; Liu et al., 2024c; Muino et al., 2024; Oyelakin et al., 2024; Tiono et al., 2024; Urbina et al., 2024), its specific function and mechanism in lavender remain poorly understood.

Herein, we used AlphaFold2 program to predict structural models of MatK1 and MatK2, and then identified potential active site residues via the GalaxyWEB program. Gene expression analysis revealed that *MatK1* was upregulated by 553.8-fold in leaves, 4.2-fold in flowers, 1.7-fold in stems, and 1.1-fold in roots at 14:00. Similarly, *MatK2* expression was upregulated by 267.5-fold in leaves, 4.2-fold in flowers, 1.3-fold in stems, and 1.0-fold in roots at 14:00. *MatK1* expression in leaves increased with the duration of heat treatment, peaking at 40°C for 3 h and 30°C for 6 h, before declining. Similarly, under salt treatment, *MatK1* expression in leaves showed a progressive increase, peaking at 300 mM NaCl for 3 h and 200 mM for 12 h, before decreasing. This study provides the first comprehensive analysis of Maturase K in *L. angustifolia*, offering valuable insights for improving the quality of lavender essential oil.

Results

Biochemical characteristics of MatK1 and MatK2

Bioinformatics analysis of the two target proteins, MatK1 and MatK2, was performed using data retrieved from the UniProt database (MatK1, entry ID A0A2R2V059; MatK2, entry ID A0A125QY04) (Figure 1, Supplementary Figure S1–S5). The molecular weights of MatK1 and MatK2 were approximately 60.31 kDa and 60.89 kDa, respectively (Table 1). Their molecular formulas were $C_{2784}H_{4317}N_{751}O_{722}S_{14}$ for MatK1 and $C_{2801}H_{4350}N_{762}O_{736}S_{13}$ for MatK2. The amino acid composition of MatK1 included 36 negatively charged residues and 70 positively charged residues, while MatK2 contained 35 negatively charged residues and 70 positively charged residues. The grand average of hydropathy (GRAVY) values for MatK1 and MatK2 were -0.10 and -0.12 , respectively (Table 1). The aliphatic indexes for MatK1 and MatK2 were 103.02 and 101.24, respectively (Table 1). Both MatK1 and MatK2 had an estimated half-life of 30 h (Table 1). The isoelectric points (pI) for MatK1 and MatK2 were 10.01

and 10.04, respectively, with protein instability indices of 51.32 and 50.08 (Table 1).

Secondary structure prediction of MatK1 and MatK2

PSIPRED analysis (Buchan et al., 2024; Jones, 1999) revealed that MatK1 contains 227 alpha helices (44.86%) in its secondary structure, along with a significant number of extended strands and random coils (Figure 2a; Table 2). Similarly, MatK2 consists of 216 alpha helices (41.86%) and numerous strands and coils in its predicted secondary structure (Figure 2b; Table 2).

Prediction and quality assessment of MatK1 and MatK2 structures

The three-dimensional (3D) structures of MatK1 and MatK2 were predicted using AlphaFold2 (Wayment-Steele et al., 2023; Jumper et al., 2021), which employs deep learning algorithms for more accurate and reliable protein structure predictions compared to traditional homology modeling methods.

To assess the quality of the predicted models (Figures 3a,d), we used the Ramachandran plot to evaluate the dihedral angles of the protein backbone, ensuring they fell within acceptable regions indicative of a stable conformation. For MatK1, 86.5% of the residues were in the most favored region, 11.9% in the additionally allowed region, 0.8% in the generously allowed region, and 0.8% in the disallowed region (Figure 3b; Table 3). For MatK2, 84.6% of residues were in the most favored region, 13.7% in the additionally allowed region, 1.4% in the generously allowed region, and 0.2% in the disallowed region (Figure 3e; Table 3).

ProSA analysis revealed Z-scores of -5.39 for MatK1 and -5.68 for MatK2 (Figures 3c,f), further supporting the high quality of the predicted models.

While the overall fold of MatK1 closely resembles that of MatK2 (Figure 4), the root mean square deviation (RMSD) for all atoms was 1.05 Å, with a sequence identity of 85.30% (Figure 4).

Predicting the active sites of MatK1 and MatK2

Using the predicted models (Figures 3–5), we utilized the GalaxyWEB program (Seok et al., 2021; Heo et al., 2013; Heo et al., 2016; Ko et al., 2012) to identify the active sites of MatK1 and MatK2. The analysis revealed that the active site residues of MatK1 are H33, N34, K51, S52, S53, and L54 (Figure 5a). For MatK2, the active site residues include K58, R59, T62, R63, and Q66 (Figure 5b). These residues are likely involved in substrate interactions, potentially forming bonds with the substrate's side chain atoms.

Gene expression profiles of *Matk1* and *Matk2* in various tissues

To investigate the spatiotemporal expression profiles of *MatK1* and *MatK2*, we performed real-time quantitative

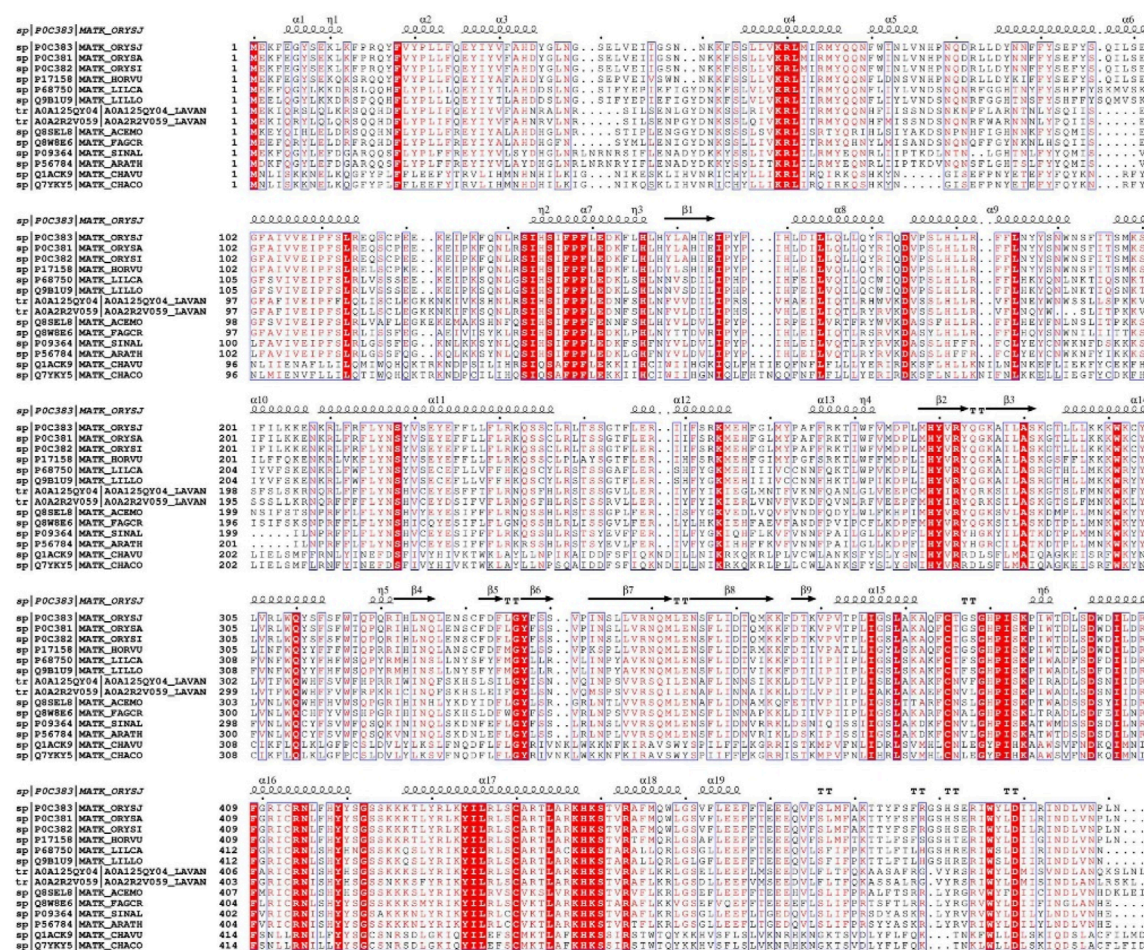


FIGURE 1

Sequence alignment of mature K family. The alignment employs the ClustalW default color scheme, where conserved amino acids are highlighted in darker shades compared to variable residues. It includes the following reference protein sequences: POC383, *Oryza sativa subsp. japonica* (Rice); POC381, *Oryza sativa* (Rice); POC382, *Oryza sativa subsp. indica* (Rice); P17158, *Hordeum vulgare* (Barley); P68750, *Lilium canadense* (Canada lily); Q9B1U9, *Lilium longiflorum* (Trumpet lily); A0A125QY04, *Lavandula angustifolia* (Lavender); A0A2R2V059, *Lavandula angustifolia* (Lavender); Q8SEL8, *Acer monspessulanum* (Montpellier maple); Q8W8E6, *Fagus crenata* (Japanese beech); P09364, *Sinapis alba* (White mustard, Brassica hirta); P56784, *Arabidopsis thaliana* (Mouse-ear cress); Q1ACK9, *Chara vulgaris* (Common stonewort); Q7YKY5, *Chara connivens* (Convergent stonewort).

TABLE 1 Characteristics of Matk1 and Matk2.

Name	Number of amino acids	Molecular weight (kDa)	Theoretical pI ^a	Instability index	Aliphatic index	GRAVY ^b	Estimated half-life (h)
Matk1	506	60.31	10.01	51.32	103.02	-0.10	30
Matk2	516	60.89	10.04	50.08	101.24	-0.12	30

Note.

^aIsoelectric point.

^bGRAVY, grand average of hydropathy.

polymerase chain reaction (RT-qPCR) using gene-specific primers (Supplementary Table S1). The results showed that the highest expression of both *Matk1* and *Matk2* occurred in the leaves at 14:00 (Figure 6). Specifically, *Matk1* expression was upregulated by 553.8-fold in leaves, 4.2-fold in flowers, 1.7-fold in stems, and

1.1-fold in roots at 14:00 (Figure 6). Similarly, *Matk2* expression was upregulated by 267.5-fold in leaves, 4.2-fold in flowers, 1.3-fold in stems, and 1.0-fold in roots at 14:00 (Figure 6). These results suggest that *Matk1* and *Matk2* are primarily involved in chloroplast photosynthesis, aligning with previous

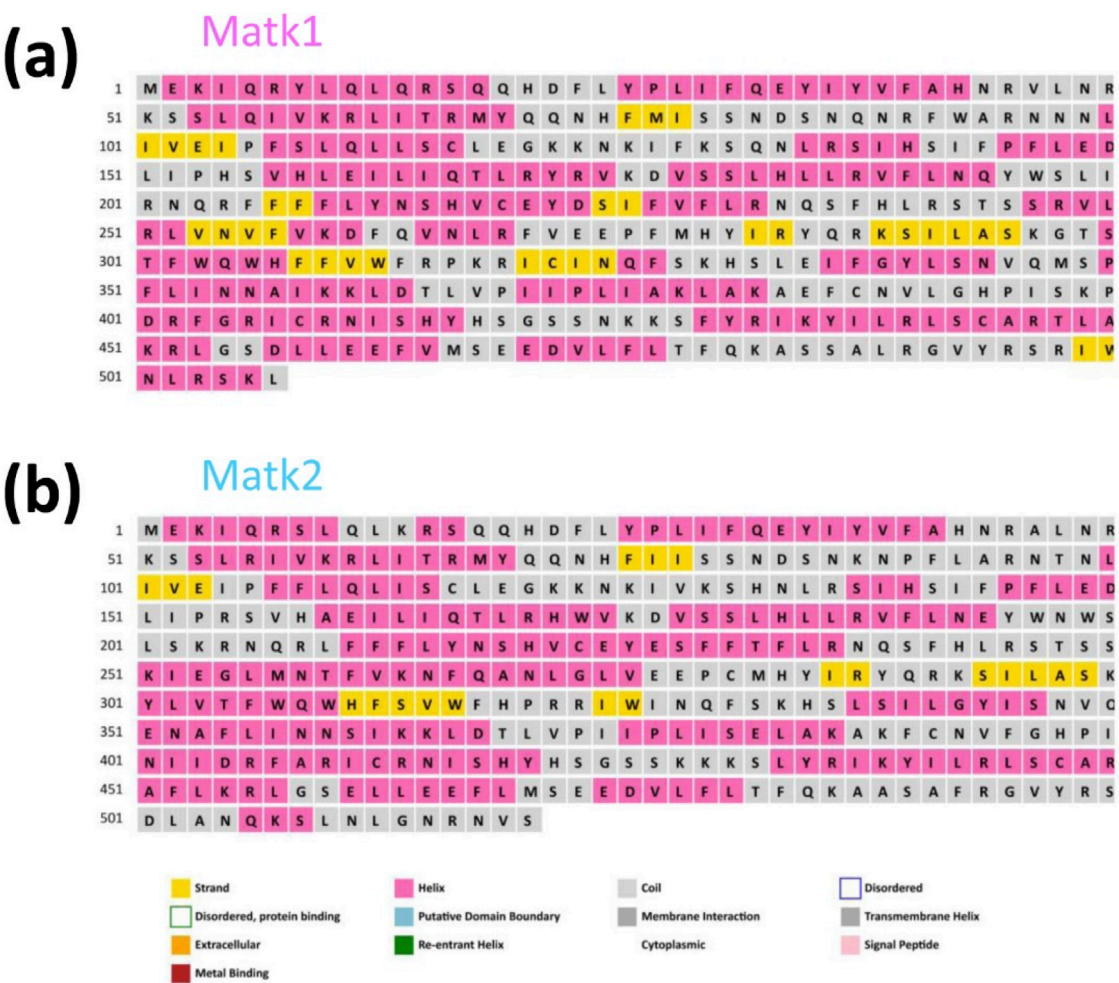


FIGURE 2 Prediction of (a) Matk1 and (b) Matk2 secondary structure models.

TABLE 2 Secondary structure prediction of Matk1 and Matk2.

Secondary structure	Alpha helix		Extended strand		Random coil	
	Number of residues	% of residues	Number of residues	% of residues	Number of residues	% of residues
Matk1	227	44.86	55	10.87	224	44.27
Matk2	216	41.86	57	11.05	243	47.09

studies (Muino et al., 2024; Hertel et al., 2013; Barthet and Hilu, 2007; Qu et al., 2018).

Expression levels of genes *Matk1* and *Matk2* under heat and salt treatments

We conducted RT-qPCR analysis to examine the expression levels of the *MatK1* gene under heat and salt treatments in leaves,

as *MatK1* exhibited higher expression in leaves compared to *MatK2* (Figure 6). The results showed that *MatK1* expression in leaves increased with the duration of heat treatment, peaking at 40°C for 3 h and 30°C for 6 h, before declining (Figure 7a). Similarly, *MatK1* expression in leaves also increased with the duration of salt treatment, peaking at 300 mM NaCl for 3 h and 200 mM for 12 h, before decreasing (Figure 7a). These findings suggested that temperature and salt concentration influence the photosynthetic rate of lavender, supporting the link between *MatK1* and photosynthesis.

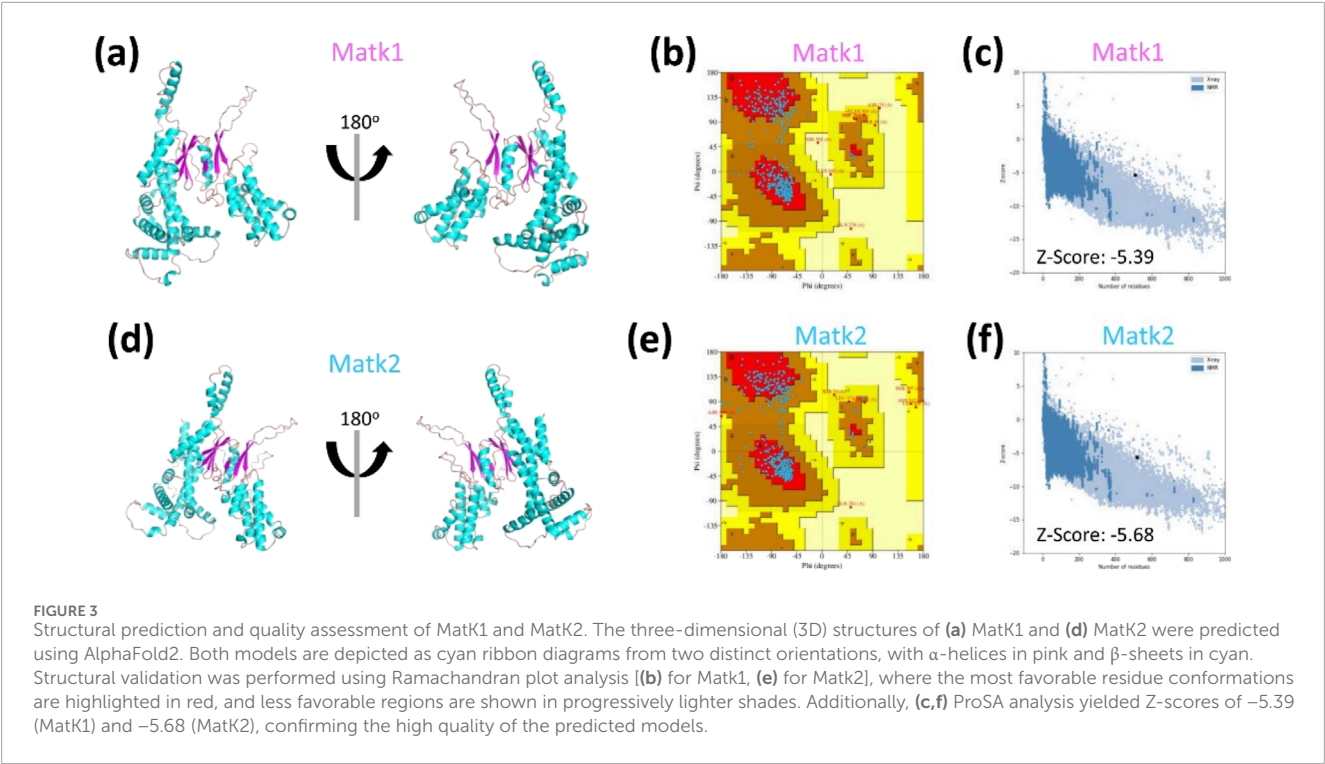


TABLE 3 Ramchandran plot analysis of structural models of MatK1 and MatK2 using PDBsum.

Residues	Residues in most favored regions		Residues in additional allowed regions		Residues in generously allowed regions		Residues in disallowed regions	
	Number of residues	% of residues ^a	Number of residues	% of residues	Number of residues	% of residues	Number of residues	% of residues
Matk1 ^a	415	86.5	57	11.9	4	0.8	4	0.8
Matk2 ^b	413	84.6	67	13.7	7	1.4	1	0.2

Note.
^aNumber of end-residues (excl. Gly and Pro): 2; Number of glycine residues: 10; Number of proline residues: 14.
^bNumber of end-residues (excl. Gly and Pro): 2; Number of glycine residues: 13; Number of proline residues: 13.

Discussion

In this work, we generated structural models using AlphaFold2 and employed the GalaxyWEB program to predict potential active site residues. At 14:00, *MatK1* expression was significantly upregulated, showing a 553.8-fold increase in leaves, 4.2-fold in flowers, 1.7-fold in stems, and 1.1-fold in roots. Similarly, *MatK2* expression increased by 267.5-fold in leaves, 4.2-fold in flowers, 1.3-fold in stems, and remained nearly unchanged (1.0-fold) in roots. Under heat stress, *MatK1* transcript levels in leaves progressively increased, peaking after 3 h at 40°C and 6 h at 30°C, followed by a decline. Similarly, under salt stress, *MatK1* expression in leaves rose with prolonged exposure, peaking after 3 h at 300 mM NaCl and after 12 h at 200 mM, before decreasing. This study provides the first comprehensive analysis of Maturase K in *L. angustifolia*, offering valuable insights that could enhance the quality of lavender essential oil.

The MatK reading frame is present in all known autotrophic land-plant chloroplast genomes containing group II introns, as well as in basal streptophyte algae (Mukhopadhyay and Hausner, 2024; Liu et al., 2018; Ho et al., 2021; Oyelakin et al., 2024). Despite their low sequence identity (Figure 1), these active sites coordinate magnesium ions (Mg²⁺), primarily via negatively charged residues. The maturase K (MatK) family may employ divergent catalytic mechanisms to promote the splicing of both its own and other chloroplast group II introns. To elucidate these mechanisms, we are examining the structural and mechanistic basis of MatK-catalyzed reactions using experimental techniques, including X-ray crystallography. In the streptophyte alga *Zygnema*, the fern *Adiantum capillus-veneris*, and the parasitic land plants *Epifagus virginiana*, *Cuscuta exaltata*, and *Cuscuta reflexa*, MatK exists as a stand-alone reading frame, with the trnK gene being absent. This suggests that MatK functions ‘in trans’, likely involved in splicing pre-RNAs other than its corresponding trnK intron

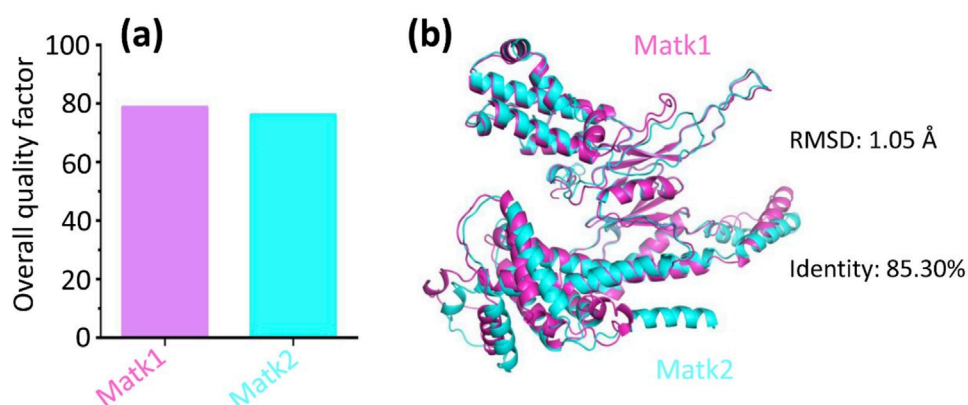


FIGURE 4

Structure comparison between Matk1 (in magenta) and Matk2 (in cyan). **(a)** The overall quality factors of structural models of Matk1 and Matk2. **(b)** Despite adopting a similar overall fold, Matk1 displayed a root mean square deviation (RMSD) of 1.05 Å (all atoms) relative to Matk2, with 85.30% amino acid sequence identity between the two proteins.

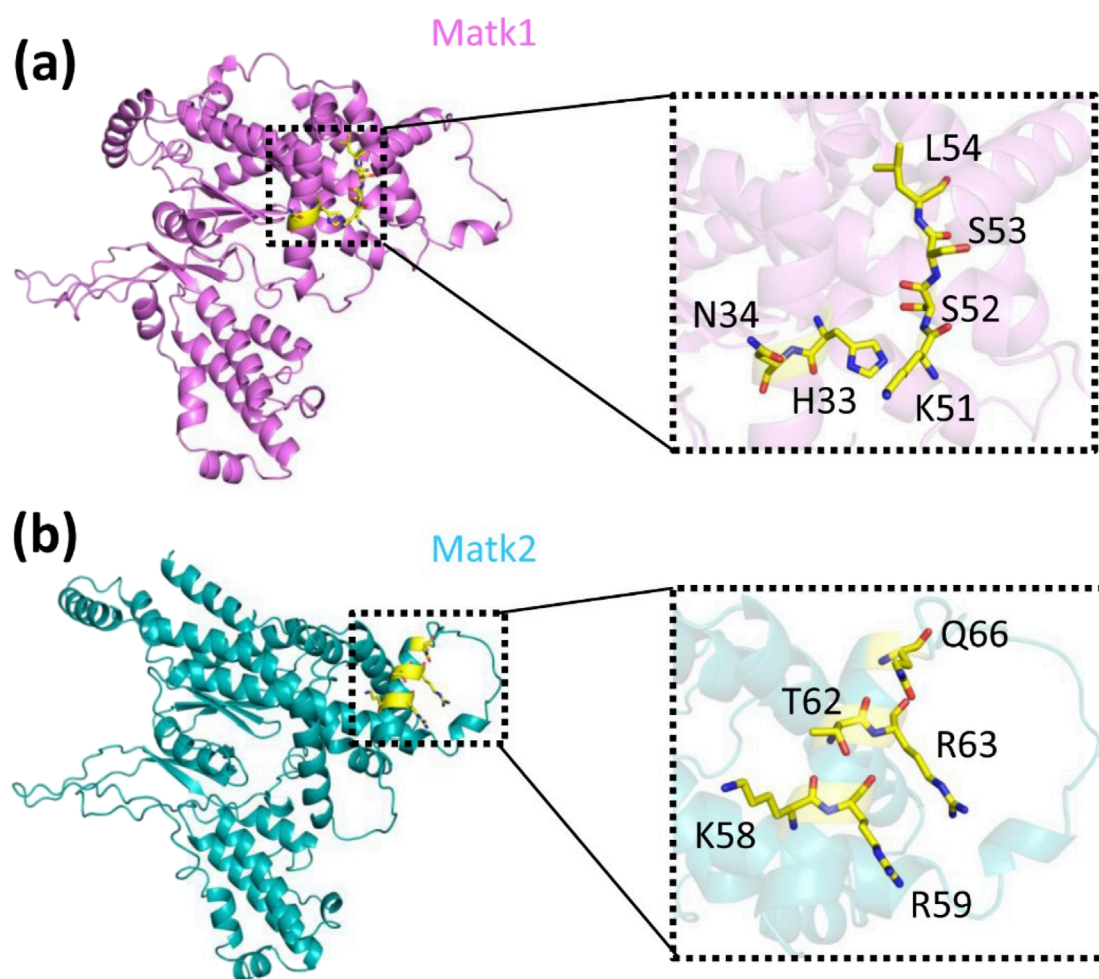


FIGURE 5

Predicting **(a)** Matk1 and **(b)** Matk2 active site residues. **(a)** The GalaxyWEB program predicted H33, N34, K51, S52, S53 and L54 as the active site residues of Matk1 (in magenta). **(b)** In Matk2 (in cyan), the active site residues identified were K58, R59, T62, R63, and Q66. The ribbon diagram of each model is shown, with a close-up view of each active site on the right.

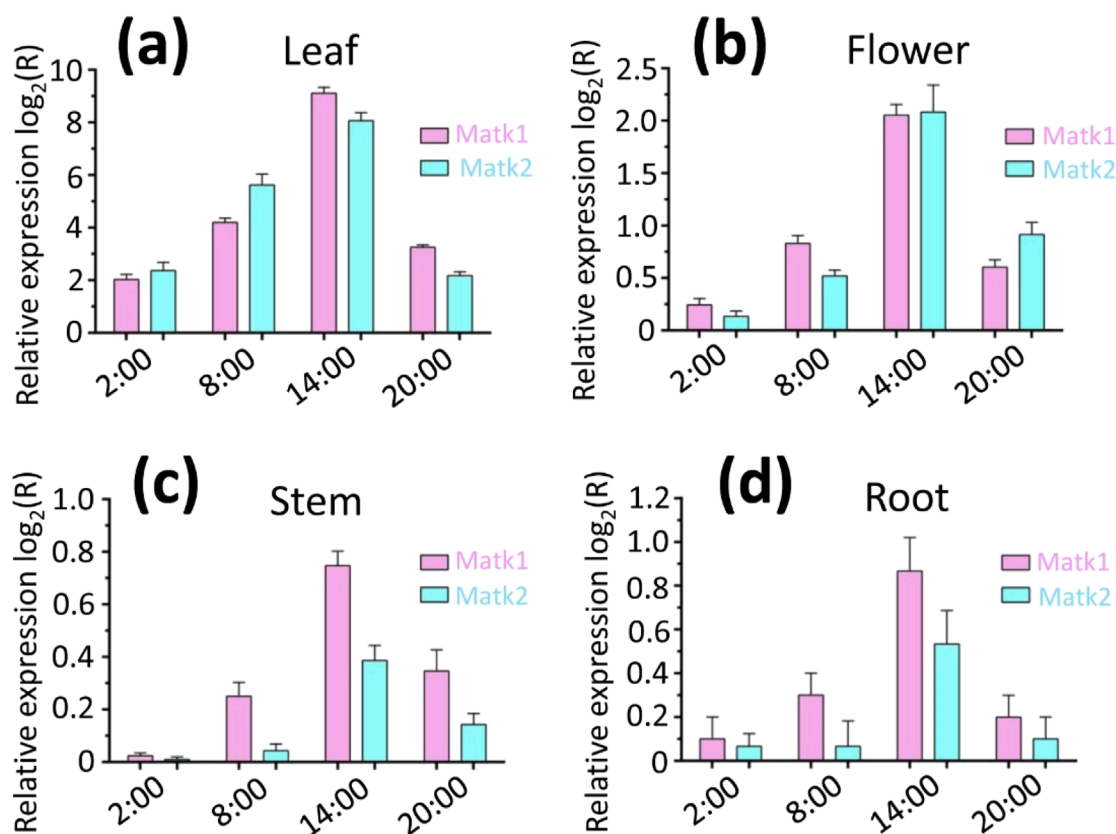


FIGURE 6

Expression levels of MatK1 and MatK2 in (a) leaf, (b) flower, (c) stem and (d) root within a 24 h day/night cycle. Relative expression analysis was conducted using RT-qPCR (real-time quantitative polymerase chain reaction). The relative expression ratios were presented as $\log_2$ values, where a ratio greater than zero indicated upregulation of gene expression.

(Hertel et al., 2013; Qu et al., 2018; Barthet and Hilu, 2007). Notably, among all analyzed embryophytes, only parasitic species have lost MatK. The retention of MatK in chloroplast genomes across early streptophytes indicates that its presence is not a random event. Furthermore, attempts at reverse genetic manipulation of the chloroplast MatK reading frame through transplastomic mutagenesis have been unsuccessful, supporting the notion that MatK is an essential gene.

To elucidate the functional role of *L. angustifolia* MatK in terpenoid biosynthesis and stress responses, we will employ a combination of *in vivo* and *in vitro* assays. Targeted knockdown of MatK via virus-induced gene silencing (VIGS) and RNA interference (RNAi) will be used to assess loss-of-function phenotypes, while *Agrobacterium*-mediated overexpression studies will evaluate gain-of-function effects on metabolic pathways. Functional validation will be further confirmed through mutant complementation in transgenic lines. These integrated approaches will systematically investigate MatK molecular mechanisms, including its potential interactions with plastid-encoded proteins and regulatory influence on secondary metabolite production. Transcriptional, translational, and metabolic changes will be monitored using quantitative PCR, Western blotting, and HPLC analyses, respectively.

In conclusion, our study offers a novel approach to comprehensively investigate the functional mechanisms of MatK

(Maturase K) in *L. angustifolia* (lavender), with the goal of enhancing the quality of lavender essential oils.

Materials and methods

Bioinformatics analysis

The amino acid sequences of MatK1 (UniProt code A0A2R2V059) and MatK2 (UniProt code A0A125QY04) (Figure 1, Supplementary Figure S1–S5) were analyzed using the ProtParam (Gasteiger, 2003; Duvaud et al., 2021) to predict their chemical properties and physicochemical parameters.

Prediction of structural models

Structural predictions of the target proteins (MatK1 and MatK2) were performed using the AlphaFold2 program (Wayment-Steele et al., 2023; Jumper et al., 2021). Secondary structures were predicted with the PSIPRED program (Jones, 1999; Buchan et al., 2024), and active site residues were identified using the GalaxyWEB program (Ko et al., 2012; Heo et al., 2013; Heo et al., 2016; Seok et al., 2021). Multiple sequence

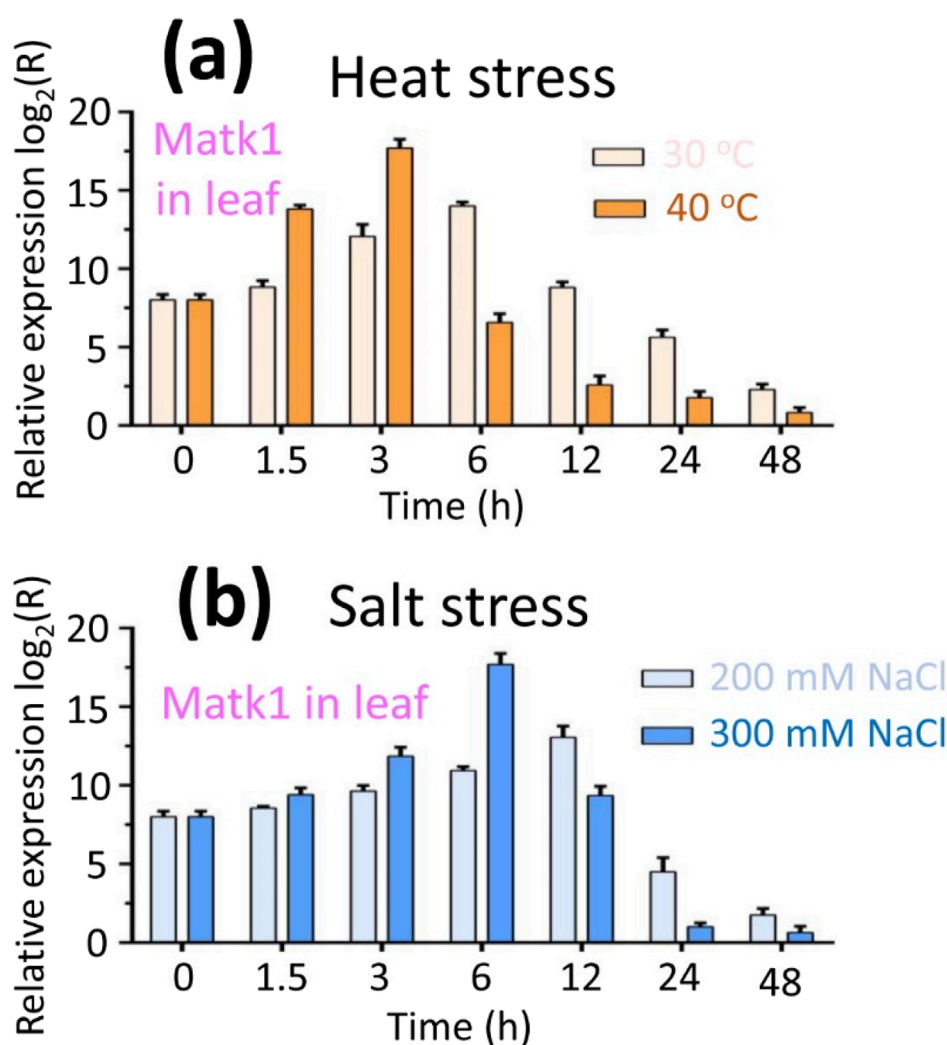


FIGURE 7 RT-qPCR data analysis of gene *MatK1* in leaf under (a) heat and (b) salt stress conditions. (a) For heat stress, plants were exposed to 30°C for 48 h and 40°C for 48 h, respectively. (b) For salt stress, plants were exposed to 200 mM NaCl for 48 h and 300 mM NaCl for 48 h, respectively. The relative expression level of the *MatK1* gene in leaf was calculated using the $2^{-\Delta\Delta CT}$ method.

alignment data were obtained from the LSQKAB program within the CCP4 suite (Collaborative Computational Project, Number, 1994), and the root mean square deviation (RMSD) for Ca atoms was calculated. Structural images were generated using PyMOL 2.3.4 (<https://www.pymol.org/2/>).

Quality assessment of structural models

To validate the tertiary structures, we used the PDBsum database (Laskowski, 2022; de Beer et al., 2014; Laskowski, 2004; 2009; Laskowski et al., 2017) to generate Ramachandran plots for *MatK1* and *MatK2*. This tool helps assess and validate protein structure quality by identifying geometric errors and improving accuracy. The Ramachandran plot specifically evaluates the stereochemical properties of the structures, displaying the dihedral angles of amino acid residues, highlighting allowed conformational regions, and identifying disallowed orientations.

Additionally, ProSA (Protein Structure Analysis) is a widely used tool for analyzing and validating predicted protein models (Wiederstein and Sippl, 2007). It aids in the analysis of protein structures derived from X-ray crystallography and NMR spectroscopy, identifying structural errors and pinpointing problematic regions, thereby improving the interpretation of the protein structures.

Expression levels of genes *MatK1* and *MatK2*

To quantify the expression levels of *MatK1* and *MatK2* under different light conditions, real-time quantitative PCR (RT-qPCR) was performed using PowerUp SYBR Green Master Mix (Applied Biosystems). Total RNA was extracted with the Universal Plant Total RNA Extraction Kit (Bioteke, Beijing, China) according to the manufacturer's instructions. cDNA was synthesized from RNA using the PrimeScript 1st Strand cDNA Synthesis Kit

(Takara, Kyoto, Japan). The primers used are listed in [Supplementary Table S1](#). RT-qPCR was conducted with the Applied Biosystems QuantStudio 5 instrument, and data were analyzed using the $2^{-\Delta\Delta CT}$ method (Green and Sambrook, 2018; Schmittgen and Livak, 2008; Livak and Schmittgen, 2001; Liu et al., 2025d; Liu et al., 2025a; Liu et al., 2025b; Liu et al., 2025c; Liu et al., 2024a; Liu et al., 2024b). Relative expression ratios are presented as $\log_2$ values in histograms. Beta-actin served as the housekeeping gene for normalization, with a positive control using the beta-actin gene. A ratio greater than zero indicated up-regulation, while a ratio less than zero indicated downregulation.

Statistical analysis

All experiments were conducted at least in triplicate. The data were expressed as mean $\pm$ SD. Statistical analysis was conducted using Origin 8.5, Microsoft Excel 2013 and SPSS 19.0. In the all statistical evaluations, $p < 0.05$ was considered statistically significant, and $p < 0.01$ was considered high statistically significant.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/[Supplementary Material](#).

Author contributions

DL: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing. NL: Investigation, Writing – original draft. DS: Investigation, Writing – original draft. ZL: Investigation, Writing – original draft.

References

- Algarni, A. A. (2022). Molecular identification and phylogenetic analysis of *Aloe shadensis* from Saudi Arabia based on matK, rbcL and ITS DNA barcode sequence. *Saudi J. Biol. Sci.* 29 (2), 1125–1133. doi:10.1016/j.sjbs.2021.09.053
- Barthet, M. M., and Hilu, K. W. (2007). Expression of matK: functional and evolutionary implications. *Am. J. Bot.* 94 (8), 1402–1412. doi:10.3732/ajb.94.8.1402
- Batiha, G. E.-S., Teibo, J. O., Wasef, L., Shaheen, H. M., Akomolafe, A. P., Teibo, T. K. A., et al. (2023). A review of the bioactive components and pharmacological properties of lavender species. *Naunyn. Schmiedeberg. Arch. Pharmacol.* 396 (5), 877–900. doi:10.1007/s00210-023-02392-x
- Braunstein, G. D., and Braunstein, E. W. (2023). Are prepubertal gynaecomastia and premature thelarche linked to topical lavender and tea tree oil use? *touchREV. Endocrinol.* 19 (2), 60–68. doi:10.17925/ee.2023.19.2.9
- Buchan, D. W. A., Moffat, L., Lau, A., Kandathil Shaun, M., and Jones David, T. (2024). Deep learning for the PSIPRED protein analysis workbench. *Nucleic Acids Res.* 52 (W1), W287–W293. doi:10.1093/nar/gkae328
- Bungau, A. F., Radu, A.-F., Bungau, S. G., Vesa, C. M., Tit, D. M., Purza, A. L., et al. (2023). Emerging insights into the applicability of essential oils in the management of acne vulgaris. *Molecules* 28 (17), 6395. doi:10.3390/molecules28176395
- Collaborative Computational Project, Number 4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* 50 (5), 760–763. doi:10.1107/s0907444994003112
- Crișan, I., Ona, A., Vârban, D., Muntean, L., Vârban, R., Stoie, A., et al. (2023). Current trends for lavender (*Lavandula angustifolia* mill.) crops and products with emphasis on essential oil quality. *Plants* 12 (2), 357. doi:10.3390/plants12020357
- de Beer, T. A. P., Berka, K., Thornton, J. M., and Laskowski, R. A. (2014). PDBsum additions. *Nucleic Acids Res.* 42 (D1), D292–D296. doi:10.1093/nar/gkt940
- de Melo Alves Silva, L. C., de Oliveira Mendes, F. C., de Castro Teixeira, F., de Lima Fernandes, T. E., Barros Ribeiro, K. R., da Silva Leal, K. C., et al. (2023). Use of *Lavandula angustifolia* essential oil as a complementary therapy in adult health care: a scoping review. *Heliyon* 9 (5), e15446. doi:10.1016/j.heliyon.2023.e15446
- Dewanjee, S., Soheli, M., Hossain, M. S., Ansari, F., Islam, M. T., Sultana, F., et al. (2023). A comprehensive review on clinically proven natural products in

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Our research work is financially supported by grants from the third batch of the “Tianchi Talent” Young Doctoral Research Grant, Xinjiang Autonomous Region (2025QNBS001), Xinjiang Key Laboratory of Lavender Conservation and Utilization (LCUZ2405), and Start-up Fund for Doctoral Research Established by Yili Normal University (2024RCYJ08).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2025.1628118/full#supplementary-material>

- the management of nerve pain, with mechanistic insights. *Heliyon* 9 (5), e15346. doi:10.1016/j.heliyon.2023.e15346
- Duvaud, S., Iobella, C., Lisacek, F., Stockinger, H., Ioannidis, V., and Durinx, C. (2021). ExPasy, the Swiss bioinformatics resource portal, as designed by its users. *Nucleic Acids Res.* 49 (W1), W216–W227. doi:10.1093/nar/gkab225
- Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31 (13), 3784–3788. doi:10.1093/nar/gkg563
- Green, M. R., and Sambrook, J. (2018). Analysis and normalization of real-time polymerase chain reaction (PCR) experimental data. *Cold. Spring Harb. Protoc.* 2018 (10). doi:10.1101/pdb.top095000
- Hedayati, S., Tarahi, M., Iraj, A., and Hashempur, M. H. (2024). Recent developments in the encapsulation of lavender essential oil. *Adv. Colloid Interface Sci.* 331, 103229. doi:10.1016/j.cis.2024.103229
- Heo, L., Lee, H., and Seok, C. (2016). GalaxyRefineComplex: refinement of protein-protein complex model structures driven by interface repacking. *Sci. Rep.* 6 (1), 32153. doi:10.1038/srep32153
- Heo, L., Park, H., and Seok, C. (2013). GalaxyRefine: protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* 41 (W1), W384–W388. doi:10.1093/nar/gkt458
- Hertel, S., Zoschke, R., Neumann, L., Qu, Y., Axmann, I. M., and Schmitz-Linneweber, C. (2013). Multiple checkpoints for the expression of the chloroplast-encoded splicing factor MatK. *Plant Physiol.* 163 (4), 1686–1698. doi:10.1104/pp.113.227579
- Ho, V. T., Tran, T. K. P., Vu, T. T. T., and Widiarsih, S. (2021). Comparison of matK and rbcL DNA barcodes for genetic classification of jewel orchid accessions in Vietnam. *J. Genet. Eng. Biotechnol.* 19 (1), 93. doi:10.1186/s43141-021-00188-1
- Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 17 (292), 195–202. doi:10.1006/jmbi.1999.3091
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596 (7873), 583–589. doi:10.1038/s41586-021-03819-2
- Khan, S. U., Hamza, B., Mir, R. H., Fatima, K., and Malik, F. (2024). Lavender plant: farming and health benefits. *Curr. Mol. Med.* 24 (6), 702–711. doi:10.2174/1566524023666230518114027
- Ko, J., Park, H., Heo, L., and Seok, C. (2012). GalaxyWEB server for protein structure prediction and refinement. *Nucleic Acids Res.* 40 (W1), W294–W297. doi:10.1093/nar/gks493
- Landmann, C., Fink, B., Festner, M., Dregus, M., Engel, K.-H., and Schwab, W. (2007). Cloning and functional characterization of three terpene synthases from lavender (*Lavandula angustifolia*). *Archives Biochem. Biophysics* 465 (2), 417–429. doi:10.1016/j.abb.2007.06.011
- Laskowski, R. A. (2009). PDBsum new things. *Nucleic Acids Res.* 37 (Database), D355–D359. doi:10.1093/nar/gkn860
- Laskowski, R. A. (2022). PDBsum1: a standalone program for generating PDBsum analyses. *Protein Sci.* 31 (12), e4473. doi:10.1002/pro.4473
- Laskowski, R. A., Chistyakov, V. V., and Thornton, J. M. (2004). PDBsum more: new summaries and analyses of the known 3D structures of proteins and nucleic acids. *Nucleic Acids Res.* 33 (Database issue), D266–D268. doi:10.1093/nar/gki001
- Laskowski, R. A., Jabłońska, J., Pravda, L., Vařeková, R. S., and Thornton, J. M. (2017). PDBsum: structural summaries of PDB entries. *Protein Sci.* 27 (1), 129–134. doi:10.1002/pro.3289
- Li, J., Zhang, X., Luan, F., Duan, J., Zou, J., Sun, J., et al. (2024). Therapeutic potential of essential oils against ulcerative colitis: a review. *J. Inflamm. Res.* 17, 3527–3549. doi:10.2147/jir.s461466
- Liu, D., Abdriyim, A., Zhang, L., and Ruzitohti, B. (2025a). Functional and mechanistic insights into the stealth protein full-length CpsY is conducive to understanding immune evasion mechanisms by *Mycobacterium tuberculosis*. *Tuberculosis* 152, 102616. doi:10.1016/j.tube.2025.102616
- Liu, D., Abdriyim, A., Zhang, L., and Yu, F. (2025b). Functional and mechanistic insights into the fatty-acid CoA ligase FadK in *Escherichia coli*. *Front. Bioscience-Landmark* 36701, 36701. doi:10.31083/FBL36701
- Liu, D., Deng, H., and Song, H. (2025c). Insights into the functional mechanisms of the sesquiterpene synthase GEAS and GERDS in lavender. *Int. J. Biol. Macromol.* 299, 140195. doi:10.1016/j.ijbiomac.2025.140195
- Liu, D., Du, Y., Abdriyim, A., Zhang, L., Song, D., Deng, H., et al. (2025d). Molecular functional mechanisms of two alcohol acetyltransferases in *lavandula x intermedia* (Lavandin). *Front. Chem.* 13, 1627286. doi:10.3389/fchem.2025.1627286
- Liu, D., Song, H., Deng, H., Abdriyim, A., Zhang, L., Jiao, Z., et al. (2024a). Insights into the functional mechanisms of three terpene synthases from *Lavandula angustifolia* (lavender). *Front. Plant Sci.* 15, 1497345. doi:10.3389/fpls.2024.1497345
- Liu, D., Tian, Z., Tusong, K., Mamat, H., and Luo, Y. (2024b). Expression, purification and characterization of CTP synthase PyrG in *Staphylococcus aureus*. *Protein Expr. Purif.* 221, 106520. doi:10.1016/j.pep.2024.106520
- Liu, X., Li, Y., Yang, H., and Zhou, B. (2018). Chloroplast genome of the folk medicine and vegetable plant *Talinum paniculatum* (jacq.) gaertn.: Gene organization, comparative and phylogenetic analysis. *Molecules* 23 (4), 857. doi:10.3390/molecules23040857
- Liu, X.-Y., Jiang, R.-C., Ma, B., Wang, Y., Yang, Y.-Z., Xu, C., et al. (2024c). Maize requires embryo defective27 for embryogenesis and seedling development. *Plant Physiol.* 195 (1), 430–445. doi:10.1093/plphys/kiae010
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25 (4), 402–408. doi:10.1006/meth.2001.1262
- Malloggi, E., Menicucci, D., Cesari, V., Frumento, S., Gemignani, A., and Bertoli, A. (2021). Lavender aromatherapy: a systematic review from essential oil quality and administration methods to cognitive enhancing effects. *Appl. Psychol. Health Well-Being* 14 (2), 663–690. doi:10.1111/aphw.12310
- Muino, J. M., Ruwe, H., Qu, Y., Maschmann, S., Chen, W., Zoschke, R., et al. (2024). MatK impacts differential chloroplast translation by limiting spliced tRNA-K(UUU) abundance. *Plant J.* 119 (6), 2737–2752. doi:10.1111/tj.16945
- Mukhopadhyay, J., and Hausner, G. (2024). Interconnected roles of fungal nuclear- and intron-encoded maturases: at the crossroads of mitochondrial intron splicing. *Biochem. Cell. Biol.* 102 (5), 351–372. doi:10.1139/bcb-2024-0046
- Oyelakin, A. S., Popoola, J. O., Babalola, F. O., Obisesan, I. A., Omotayo, O. E., Oluwatuyi, V. O., et al. (2024). Dataset on MatK-based intra- and inter-specific genetic relationships among four solanum L. species from Southwestern Nigeria. *Data Brief* 56, 110815. doi:10.1016/j.dib.2024.110815
- Qu, Y., Legen, J., Arndt, J., Henkel, S., Hoppe, G., Thieme, C., et al. (2018). Ectopic transplastomic expression of a synthetic MatK gene leads to cotyledon-specific leaf variegation. *Front. Plant Sci.* 9, 1453. doi:10.3389/fpls.2018.01453
- Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3 (6), 1101–1108. doi:10.1038/nprot.2008.73
- Seok, C., Baek, M., Steinegger, M., Park, H., Lee, G. R., and Won, J. (2021). Accurate protein structure prediction: what comes next? *Biodesign* 9 (3), 47–50. doi:10.34184/kssb.2021.9.3.47
- Tiono, Y. V., Prasetyo, A. H., Wulanjati, M. P., Wink, M., and Nurcahyanti, A. D. R. (2024). Inhibition of oxidative stress of *Biancaea sappan* (L) tod. From java. *Nat. Prod. Res.*, 1–7. doi:10.1080/14786419.2024.2355585
- Tripodi, P. (2023). Application of high-resolution melting and DNA barcoding for discrimination and taxonomy definition of rocket salad (*diplotaxis* spp.) species. *Genes* 14 (8), 1594. doi:10.3390/genes14081594
- Unnikrishnan, R., Sumod, M., Jayaraj, R., Sujanal, P., and Dev, S. A. (2021). The efficacy of machine learning algorithm for raw drug authentication in *Coscinium fenestratum* (gaertn.) colebr. Employing a DNA barcode database. *Physiology Mol. Biol. Plants* 27 (3), 605–617. doi:10.1007/s12298-021-00965-9
- Urbina, H., Jones, C., Moore, M. R., and Gazis, R. (2024). Susceptibility of centipede tongavine, *Epipremnum pinnatum*, commercially grown in nurseries in Florida to aroid leaf rust, *Pseudocercospora paullula*. *Plant Dis.* 108 (1), 217. doi:10.1094/pdis-07-23-1360-pdn
- Wayment-Steele, H. K., Ojoawo, A., Otten, R., Apitz, J. M., Pitsawong, W., Hömberger, M., et al. (2023). Predicting multiple conformations via sequence clustering and AlphaFold2. *Nature* 625 (7996), 832–839. doi:10.1038/s41586-023-06832-9
- Wiederstein, M., and Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35 (Web Server), W407–W410. doi:10.1093/nar/gkm290
- Zhang, D., Xu, H., Gao, J., Portieles, R., Du, L., Gao, X., et al. (2021). Endophytic *Bacillus altitudinis* strain uses different novelty molecular pathways to enhance plant growth. *Front. Microbiol.* 12, 692313. doi:10.3389/fmicb.2021.692313