



Utilizing evolutionary and structural evidence to identify key regulatory elements in genistein biosynthesis in *Medicago truncatula*

Hadiseh Sadat Hosseini Pouya^{1*}, Parviz Heidari², Monireh Cheniany¹

¹ Department of Biology, Faculty of Science, Ferdowsi University of Mashhad, Iran, *hosseinipouya@mail.um.ac.ir

² Department of Agronomy and Plant Breeding, Faculty of Agriculture, Shahrood University of Technology, Iran

Abstract

Genistein is an important isoflavone in legumes (including the model species *Medicago truncatula*). This metabolite has notable medicinal and agricultural importance. The present study aims to identify key evolutionary and structural features and to analyze the cis-regulatory elements of the *HID* and *IFS* gene families, which encode two essential enzymes in the genistein biosynthesis pathway. For this purpose, *in silico* analyses were performed using *M. truncatula* genome, and the evolutionary relationships and protein structures of *HID* genes were also examined. In addition, promoter sequences of all *MtHID* and *MtIFS* genes were analyzed. The results showed that the *HID* protein structure is conserved across legumes, and promoter analysis revealed that stress-responsive elements, particularly MYB and MYC, are the most abundant regulatory motifs in both gene families. Additionally, ABA-responsive elements were the most abundant in *MtHID* promoters, while both ABA- and MeJA-responsive elements were dominant in *MtIFS* promoters. Based on these results, the cis-elements suggest a multi-layered regulatory structure in the genistein biosynthesis pathway. These findings provide valuable insights for future approaches to enhance genistein production through targeted manipulation of regulatory networks and metabolic engineering.

Introduction

Medicago truncatula is an important model plant used to study molecular, genetic, and physiological pathways in legume crops (Zhang *et al.*, 2022). In addition to its importance in agricultural research, *M. truncatula* also produces valuable secondary metabolites, including genistein. Genistein has been shown to exhibit protective effects against a range of human diseases, including osteoporosis, cancer, diabetes, and cardiovascular disorders (Zaheer and Akhtar, 2017). In addition, this metabolite plays an important role in defense against biotic stresses, acts as an antioxidant under abiotic stresses, and enhances the symbiotic relationships of legumes (Trush and Pal'ove-Balang, 2023). The enzymes involved in the genistein biosynthesis pathway that catalyze the conversion of naringenin to genistein in three steps include isoflavone synthase (*IFS*), cytochrome P450 reductase (*CPR*), and 2-hydroxyisoflavone dehydratase (*HID*) (Kim *et al.*, 2024). In this study, key enzymes involved in the genistein biosynthesis pathway, including *IFS* and *HID*, were analyzed. The evolutionary relationships of *HID* genes and their protein structure were examined. Then, promoter analysis of each member of the *IFS* and *HID* gene families was performed to identify potential regulatory elements influencing their expression. These findings provide a basis for designing strategies to enhance genistein production in plants.

Materials and methods

The *HID* and *IFS* gene families were identified in the model legume *M. truncatula* using the Ensembl Plants database. Subsequently, the protein sequences of identified genes were retrieved, and their 5' flanking regions were extracted from the Ensembl Plants database to examine the promoter sequences. Then, the protein sequences of *HID* homologs from *M. truncatula*, *Glycine max*, and *Glycine soja* were aligned using Clustal Omega. The phylogenetic tree was constructed using iTOL. Protein domains, including α/β -hydrolase fold-3 in *HID* proteins and Cytochrome P450 in *IFS* proteins, were analyzed to confirm gene identity using Pfam and CDD tools. The position of the conserved α/β -hydrolase fold domain in the *HID* protein sequences was mapped using TBtools for the three species mentioned above.

Analysis of Upstream regions of *HID* and *IFS* genes was performed using PlantCARE to identify regulatory elements potentially involved in genistein biosynthesis. TBtools, GraphPad Prism 9.0, and Python (v3.x) were used to plot and visualize the results.

Results and discussion

Analysis of the evolutionary relationships and structures of *HID* and *IFS* proteins revealed that these enzymes are highly conserved in structure and function. Promoter analysis suggests that stress is the primary factor regulating the expression of their gene families. Hormones and light are key regulators of both gene families. In *MtHIDs*, ABA-responsive elements were the most abundant, whereas in *MtIFS* genes, ABA- and MeJA-responsive elements were the most frequent. Differences in the composition of regulatory elements, such as the absence of gibberellin-responsive cis-elements in *MtIFSs* compared to *MtHID* genes, may indicate distinct regulatory mechanisms and suggest a multi-layered regulatory architecture for the genistein biosynthesis pathway. Based on the findings of this study, it is recommended that stress treatments and relevant plant hormones be applied to evaluate their effects on genistein accumulation in legumes.

References

- Kim, J.M., Lee, J.W., Seo, J.S., and Ha, B. (2024). Differentially Expressed Genes Related to Isoflavone Biosynthesis in a Soybean Mutant Revealed by a Comparative Transcriptomic Analysis. *Plants*, 13(5):584. doi: <https://doi.org/10.3390/plants13050584>
- Trush, K., and Pal'ove-Balang, P. (2023). Plant Stress Biosynthesis and role of isoflavonoids in legumes under different environmental conditions. *Plant Stress*, 8: 100153. doi: <https://doi.org/10.1016/j.stress.2023.100153>
- Zaheer, K., and Akhtar, M.H. (2017). An updated review of dietary isoflavones: Nutrition, processing, bioavailability and impacts on human health. *Critical Reviews in Food Science and Nutrition*, 57(6):1280–1293. doi: <https://doi.org/10.1080/10408398.2014.989958>
- Zhang, X., Sun, Y., Qiu, X., Lu, H., and Hwang, I. (2022). Tolerant mechanism of model legume plant *Medicago truncatula* to drought, salt, and cold stresses. *Frontiers in Plant Science*, 13:847166. doi: <https://doi.org/10.3389/fpls.2022.847166>