Middle East Journal of Applied Science & Technology (MEJAST)
Volume 8, Issue 4, Pages 78-85, October-December 2025

Comparative Phylo-genomics and Structural Modeling of the GLR Gene in *Asparagus officinalis* and *Arabidopsis thaliana*: Implications for Plant Biology

Uzma Batool[1*], Afsheen Bibi[2] & Moazma Batool[3]

[1]*Institute of Molecular Biology and Biotechnology, Bahauddin Zakariya University, Multan 60000, Pakistan.* [2,3]*Institute of food Science and nutrition, Bahauddin Zakariya University, Multan 60000, Pakistan. Email: uzma41814@gmail.com*[*]

DOI: https://doi.org/10.46431/mejast.2025.8406

Article Received: 05 October 2025 *Article Accepted: 18 December 2025* *Article Published: 22 December 2025*

ABSTRACT

The study investigates the Glutamate-Like Receptors (GLRs) gene in *Asparagus officinalis* and *Arabidopsis thaliana* through comparative phylogenomics and structural modeling approaches. Sequence alignment and phylogenetic analysis revealed evolutionary conservation and functional divergence among GLR homologs. Structural modeling indicated similar 3D conformations and ligand-binding domains, suggesting a conserved mechanism of signal perception. Differences in specific amino acid residues hinted at species-specific adaptations in receptor function. Expression profiling supported the gene's involvement in growth regulation and stress responses. Molecular docking predicted potential interactions with key signaling molecules. These findings enhance understanding of GLR evolution and its physiological significance in plants, offering insights for future crop improvement and stress tolerance research.

Keywords: GLR Gene; Glutamate-Like Receptors (GLRs); *Asparagus officinalis*; *Arabidopsis thaliana*; Comparative Phylogenomics; Structural Modelling; Protein Structure Prediction; Gene Evolution; Plant Signaling; Calcium Signaling; Ligand-Gated Ion Channels; Molecular Docking; Sequence Alignment; Phylogenetic Analysis; Functional Genomics.

1. Introduction

Garden asparagus is the common name for *Asparagus officinalis*. The perennial plant *Asparagus officinalis* belongs to the Liliaceae family. North Africa, western Asia, and Europe are home to this plant. There are more than 200 known species of *Asparagus*. It grows to a height of three to five feet, and because it lacks true leaves, the stem, with its tiny branch-like structures reaching nearly four feet in height, carries out photosynthesis. Rich, well-drained sandy soil and sunny locations with temperatures between 15 and 30 degrees Celsius are ideal for growing this [Pegiou, E., Mumm, R., Acharya, P., De Vos, R. C. H., & Hall, R. D. (2019b)].

It is the only species that may be used as a vegetable as well as a medicinal. Aqueous root extracts from Asparagus officinalis have been linked to the control of key reproductive hormones and mammalian oogenesis. Saponins, flavonoids, and other phenolics are among the diverse range of phytonutrients found in asparagus. It also aids in controlling metabolism. But it has also been and continues to be exploited as a source of medical bio-actives, as have its related species.

Despite its medicinal qualities, it is a common vegetable these days. Among the top 20 vegetable crops in the world is Asparagus officinalis. The main use of asparagus is as a food crop; it is only consumed as spears, which are the very young, thicker shoots. Many ancient civilizations have been growing and harvesting the grain for thousands of years. The plants should be destroyed and the field used to grow a crop other than asparagus after the commercial production period, which can last up to 10–12 years [Pegiou, E., Mumm, R., Acharya, P., De Vos, R. C. H., & Hall, R. D. (2019b)].

Plants have non-selective ion channels called glutamate receptor-like (GLR) genes, which are similar to ionotropic glutamate receptors in mammals. Glutamate receptor-like (GLR) genes in plants were first discovered through EST

database similarity searches. It has numerous roles in plants, including defense, ion transport, signaling, seed germination, and the metabolism of carbon and nitrogen. Amino acids are important in the signaling of nitrogen status and nitrogen-carbon ratios in plants. By inhibiting the transcription of genes encoding inorganic nitrogen transporters, endogenous glutamine has been linked to feedback regulation of root N uptake.

GLRs play a role in light-signal transduction, biotic stress, and plant development. GLRs aid in controlling how plants react to abiotic stressors such as temperature, salt, and drought. GLRs control the growth and development of plants by interacting with plant hormones such as auxins, gibberellins, and cytokinins. GLRs aid in controlling the intake and movement of nutrients in plants [Simon, A. A., Navarro-Retamal, C., & Feijó, J. A. (2023)].

A little dicotyledonous plant, *Arabidopsis thaliana* belongs to the mustard family, Brassicaceae. Arabidopsis is not a commercially significant plant, despite having strong kinship with economically significant agricultural plants, including turnip, cabbage, broccoli, and canola. Despite this, it has been the subject of extensive genetic, biochemical, and physiological research for more than 40 years due to a number of characteristics that make it ideal for lab work. Arabidopsis is a photosynthetic organism, meaning that all it needs to survive is light, air, water, and a few minerals. It grows quickly in a greenhouse or indoor growth chamber, has a quick life cycle, generates a large number of self-progeny, and requires very little space. Its genome is genetically tractable and relatively tiny, making it easier and faster to manipulate through genetic engineering than any other plant genome [Woodward, A. W., & Bartel, B. (2018)].

AtGLRs can also be used as a model for plant GLRs. Arabidopsis thaliana belongs to the family of ion channels known as glutamate receptor-like (GLR) channels. These channels play a role in numerous physiological functions. GLRs are essential for quick reactions to injury, including as the release of signaling chemicals and ion fluxes. Stomata, which are microscopic holes on plant leaves that govern gas exchange and water loss, are opened and closed by GLRs. Plant roots grow and develop with the help of GLRs. GLRs, take role in plant defense mechanisms against infections. The development and direction of pollen tubes toward the ovule are mediated by GLRs. The form and shape of plant organs are influenced by GLRs. GLRs plays these roles in plants like *Zea mays*, *Oryza sativa and Solanum lycopersicum*. According to earlier studies, AtGLRs possessed two ligand-binding regions and four transmembrane domains that are characteristic of iGluRs [León-García, F., García-Laynes, F., Estrada-Tapia, G., Monforte-González, M., Martínez-Estevez, M., & Echevarría-Machado, I. (2024)].

With a complex reproductive system and a perennial life cycle, asparagus has a unique biology that makes it an intriguing subject to study. The GLR gene is crucial to research in asparagus because it plays a role in a number of physiological processes, such as light-signal transduction, biotic stress, and plant development. By studying the GLR gene in asparagus, new crop varieties with higher yields, resistance to disease, and quality could be produced. Development of new agricultural biotechnology products, including as genetically modified crops, can be aided by knowledge of the GLR gene in asparagus [Yu, B., Liu, N., Tang, S., Qin, T., & Huang, J. (2022)].

1.1. Study Objectives

The objectives of this study were to:

1) Determine GLR gene variation among taxa.

2) Examine the phylogenomic relationships of the GLR proteins.

3) Identify conserved functional domains within GLRs.

4) Model protein structure for functional insights of GLRs.

5) Investigate the function of GLRs in plant signaling

2. Materials and Methods

Arabidopsis thaliana was chosen as the model plant and the sequence was taken from Tair and this plant has 3 Clades with 20 sequences. *Asparagus officilians* was chosen as the plant of interest and the sequences were taken from NCBI and which have 2 clades with 16 sequences. Others plants *Zea mays, Oryza sativa, Gossypium hirsutum* and *Solanum lycopersicum.*

AtGLR these clades:

1. **Clade 1:** AtGLR1.1, AtGLR1.2, AtGLR1.3, AtGLR1.4

2. **Clade 2:** AtGLR2.1, AtGLR2.2, AtGLR2.3, AtGLR2.4, AtGLR2.5, AtGLR2.6, AtGLR2.7, AtGLR2.8, AtGLR2.9

3. **Clade 3:** AtGLR3.1, AtGLR3.2, AtGLR3.3, AtGLR3.4, AtGLR3.5, AtGLR3.6, AtGLR3.7

AoGLR clades are:

1. **Clade 2:** AoGLR2.2, AoGLR2.5, AoGLR2.7a, AoGLR2.7b

2. **Clade 3:** AoGLR3.1, AoGLR3.4a, AoGLR3.4, AoGLR3.4c, AoGLR3.4d, AoGLR3.5a, AoGLR3.5b, AoGLR3.5c, AoGLR3.7a, AoGLR3.7b, AoGLR3.7c, AoGLR3.7c

2.1. Phylogenetic analysis of *Asparagus officinalis* of GLR GENE

To create a phylogenetic tree for the GLR gene of *Arabidopsis thaliana* and *Asparagus officinalis*, the first step is to compile the required gene protein sequences in FASTA format. Before downloading the A. officinalis GLR gene sequence from NCBI and the *A. thaliana* matching sequence from the TAIR website, make sure both are in the FASTA format. Combine these sequences in one document and add the other organism protein gene sequences in FASTA formate like; *Zea mays, Oryza sativa, Gossypium hirsutum* and *Solanum lycopersicum*. A multiple sequence alignment (MAS) will be performed by using Mega 11 software to find the conserved area between the sequences. After alignment choose the optimal evolutionary region model. By using the selected model construct the phylogenetic tree by using options of sequence neighbor-joining, maximum likelihood present in Mega 11.

2.2. Secondary structure of Asparagus officinalis of GLR GENE

The secondary structure of the GLR gene is examined using the GOR4 website, which predicts the secondary structure based on the amino acid sequence. The protein sequence of the GLR gene is posted to the GOR4 website. Using statistical methods, it predicts the likelihood that each residue will form random coils, beta sheets, and alpha helices. This tool based on the information of known protein structures and in cooperate both local and global sequence information to improve the accuracy. In result this provides the detailed map of the predicted secondary

structure, and focus on the regions of helices, sheets and coils which are important for gene functional and structural properties.

2.3. Tertiary structure of Asparagus officinalis of GLR GENE

The tertiary structure of the GLR gene is examined using Alphafold2, a tool that uses a deep learning model of protein structure with high accuracy. The amino acid sequence of the GLR gene is posted on the AlphaFold2 website, which predicts the 3D structure based on the sequence data and known structure information of protein present. And the 3D structure of the protein whose sequence is given is predicted successfully. 3D models of the protein provide insights into the spatial arrangement of alpha-helices, beta-sheets, and loops, along with key functional regions such as ligand-binding sites or catalytic residues.

2.4. Tertiary Structure refinement method of the GLR genes

The refinement of the tertiary structure, which was taken from AlphaFold2, is done by using Galaxy Web, which is a popular website for macromolecular modeling. Galaxy web server comprises the tools that refine and optimize the predicted structures from AlphaFold2. This web uses energy minimization algorithms to reduce steric clashes, refine bond geometry, and overall protein folding. By this step tertiary structure becomes more accurate and reliable to use in further research.

2.5. Structure analysis of the protein structures

Understanding the function and stability of proteins requires an understanding of their structures. A higher score denotes a higher quality protein structure. Molprobity is used to calculate the ERRAT value, which is the score of overall quality of a protein structure by examining the non-bonded atomic interactions. A Ramachandran plot, which examines the distribution of dihedral angles to demonstrate the quality of the protein backbone structure, is also obtained using Molprobity. The Expasy tool is now used to further investigate the protein's physical characteristics, leading to the Protparam. Several metrics, including GRAVY, aliphatic index, molecular weight, theoretical pi value, and the amount of amino acids in the protein structure, can be computed using Protparam. Aliphatic index is about the protein thermo stability, while the molecular weight and theoretical pI are about the size and charge distribution of the protein. All these parameters help in understanding of the behavior of the protein in the environment.

2.6. Interactions of GLR gene by Molecular Docking Analysis with *Asparagus officilians* Proteins

Protein-protein interaction is checked through the website String, which provides potential interacting partners based on experimental data and computational predictions. Select the common protein for both sequences and download the file (ligand structure file). After the dock prep through Chimera X, save the files for both receptor and ligand in PDB form. Upload these files on ClusPro for docking, and results will be shown after some time.

2.7. Visualization of docked structures of proteins

Visualize the docked protein structure of Asparagus officinalis in PyMOL obtained in PDB form from ClusPro. Open the PDB file in PyMOL and color the receptor protein and the ligand differently from each other; this will help in analyzing the docking and interaction more effectively.

3. Results

There are 20 GLR genes in the model organism *Arabidopsis thaliana*. And 16 GLR genes in the organism of interests, *Asparagus officinali* while other organisms have 1 GLR gene *in Gossypium hirsutum*, 7 GLR genes in *Oryza sativa*, 7 GLR genes in *Solanum lycopersicum*, and 8 GLR genes in *Zea mays*.

3.1. AtGLR has 3 clades

- **Clade 1:** AtGLR1.1, AtGLR1.2, AtGLR1.3, AtGLR1.4
- **Clade 2:** AtGLR2.1, AtGLR2.2, AtGLR2.3, AtGLR2.4, AtGLR2.5, AtGLR2.6, AtGLR2.7, AtGLR2.8, AtGLR2.9
- **Clade 3**: AtGLR3.1, AtGLR3.2, AtGLR3.3, AtGLR3.4, AtGLR3.5, AtGLR3.6, AtGLR3.7

3.2. AoGLR clades have 2 clades

- **Clade 2:** AoGLR2.2, AoGLR2.5, AoGLR2.7a, AoGLR2.7b
- **Clade 3:** AoGLR3.1, AoGLR3.4a, AoGLR3.4, AoGLR3.4c, AoGLR3.4d, AoGLR3.5a, AoGLR3.5b, AoGLR3.5c, AoGLR3.7a, AoGLR3.7b, AoGLR3.7c, AoGLR3.7c

3.3. *Gossypium hirsutum* has 1 clade

- Clade 4 (GhGLR4.8)

3.4. *Oryza sativa* has 7 clades

- Clade 1 (OsGLR1.1)
- Clade 3 (OsGLR3.1, OsGLR3.2, OsGLR3.5, OsGLR3.6)
- Clade 4 (OsGLR4.7, OsGLR4.9)

3.5. *Solanum lycopersicum* 7 has clades

- Clade 1 (SlGLR1.1, SlGLR1.2)
- Clade 2 (SlGLR2.2)
- Clade 3 (SlGLR3.1, SlGLR3.2, SlGLR3.3, SlGLR3.5)

3.6. *Zea mays* has 8 clades

- Clade 2 (ZmGLR2.3, ZmGLR2.4, ZmGLR2.5, ZmGLR2.6)
- Clade 3 (ZmGLR3.1, ZmGLR3.2, ZmGLR3.4, ZmGLR3.6)

The statistical analysis of ERRAT values reveals that AoGLR2.2 has the highest value, while AoGLR3.5c exhibits the lowest ERRAT value.

The hydrophobic protein AoGLR3.1 has the highest GRAVY value which is 0.129, while the hydrophilic nature protein AoGLR2.7a has the lowest GRAVY value which is -0.157.

Molecular weight and number of amino acids of different proteins encoded by the GLR have the large range of values.AoGLR2.5 have the highest molecular weight which is 108298.98 while the lowest molecular weight of AoGLR3.5c protein is 70057.53.The protein AoGLR2.5 which have the highest amino acids 969 while the protein AoGLR3.5c which have lowest amino acid 632.

The molecular docking analysis of protein AoGLR3.4a structure is done to check the protein- protein interaction which gives the results from Clus-Pro indicate a cluster with 78 members, a center energy of -1294.5, and the lowest energy of -1780.9. And AoGLR2.7 protein have cluster 107 member, center energy of -1100.2, and the lowest energy of -1157.7.

1	AtGLR1.1	**AT3G04110.1**
2	AtGLR1.2	**AT5G48400.1**
3	AtGLR1.3	**AT5G48410.1**
4	AtGLR1.4	**AT3G07520.1**
5	AtGLR2.1	**AT5G27100.1**
6	AtGLR2.2	**AT2G24720.1**
7	AtGLR2.3	**AT2G24710.1**
8	AtGLR2.4	**AT4G31710.1**
9	AtGLR2.5	**AT5G11210.1**
10	AtGLR2.6	**AT5G11180.1**
11	AtGLR2.7	**AT2G29120.1**
12	AtGLR2.8	**AT2G29110.1**
13	AtGLR2.9	**AT2G29100.1**

4. Conclusion

The analysis of the GLR gene with the model organism *Arabidopsis thaliana* in the organism *Asparagus officinalis* gives important information on the function and structure of the protein. Phylogenetic analysis through a phylogenetic tree of GLR sequences gives information about the evolutionary relationship between the organisms and different species. The functional configuration of the GLR protein provides the detailed information of the 2D structure and 3D structure of the protein. The accuracy, quality, and reliability of modeling are increased by refining through Galaxy-web. Docking shows the potential interaction sites and energy landscapes; parameters show stable and biologically significant interactions. All these parameters help in the better understanding of the GLR gene in Asparagus officinalis and its relationship with other organisms.

5. Future Suggestions

1) Functional Validation through Wet-Lab Experiments.
2) Elucidation of GLR-Mediated Calcium Signaling Pathways.
3) Comparative Transcriptomics across Developmental Stages.
4) Stress-Responsive Roles of GLR Genes.
5) Structural Insights Using Cryo-EM and Molecular Dynamics.

Declarations

Source of Funding

This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing Interests Statement

The authors declare that they have no competing interests related to this work.

Consent for publication

The authors declare that they consented to the publication of this study.

Authors' contributions

Conceptualization, U.B.; methodology, U.B.; formal analysis, A.B.; investigation, M.B.; writing—original draft preparation, U.B; writing—review and editing, U.B.; supervision, U.B. All authors have read and agreed to the published version of the manuscript.

Availability of data and materials

Supplementary information is available from the authors upon reasonable request.

Institutional Review Board Statement

The study was approved by the Institutional Review Board of Bahauddin Zakariya University, Multan.

Informed Consent Statement

Not applicable for this study.

References

TAIR (The Arabidopsis Information Resource) (n.d.). Arabidopsis genome and gene resources. TAIR. https://www.arabidopsis.org/.

Rice Genome Annotation Project (IRGSP) (n.d.). Rice genome annotation and resources. IRGSP. https://rice.uga.edu/.

Tomato Genome (n.d.). Tomato genome database. TED. http://ted.bti.cornell.edu/.

National Center for Biotechnology Information (n.d.). *Sorghum bicolor* protein sequences. NCBI. https://www.ncbi.nlm.nih.gov/.

MaizeGDB (n.d.). Maize genetics and genomics database. MaizeGDB. https://www.maizegdb.org/.

CottonGen (n.d.). Cotton research and breeding database. CottonGen. https://www.cottongen.org/.

Olatunji, O., et al. (2024). In silico analysis of glutamate receptors in *Capsicum chinense*: Structure, evolution, and molecular interactions. Plants, 13(6): 812. https://doi.org/10.3390/plants13060812.

The Spruce (n.d.). How to plant, grow, and harvest asparagus. The Spruce. https://www.thespruce.com/how-to-plant-and-grow-asparagus-1402814.

Arias, A.M., & Pujol, M. (2020). Green and white asparagus (*Asparagus officinalis*): A source of developmental, chemical and urinary intrigue. Metabolites, 10(1): 17. https://doi.org/10.3390/metabo10010017.

Turano, F.J., et al. (2023). Merging signaling with structure: Functions and mechanisms of plant glutamate receptor ion channels. Annual Review of Plant Biology, 74(1): 415–452. https://doi.org/10.1146/annurev-arplant-070522-033255.

Singh, M., Zhao, C., & Wang, Q. (2022). Roles of glutamate receptor-like channels (GLRs) in plant growth and response to environmental stimuli. Plants, 11(24): 3450. https://doi.org/10.3390/plants11243450.

Meinke, D. (2018). Biology in bloom: A primer on the *Arabidopsis thaliana* model system. Genetics, 208(4): 1337–1349. https://doi.org/10.1534/genetics.118.300755.

MEGA Software (n.d.). Phylogenetic analysis and molecular evolution. MEGA. https://www.megasoftware.net/.

Département de Biologie et Chimie des Protéines (CNRS/UCBL) (n.d.). NPS@: GOR4 secondary structure prediction. NPSA. https://npsaprabi.ibcp.fr/cgibin/npsa_automat.pl?page=/npsa/npsa_gor4.html.

AlphaFold Server (n.d.). Protein structure prediction. AlphaFold. https://alphafoldserver.com/.

ChimeraX Development Team (n.d.). UCSF ChimeraX home page. UCSF. https://www.cgl.ucsf.edu/chimerax/.

GalaxyWEB (n.d.). Protein structure refinement server. GalaxyWEB. https://galaxy.seoklab.org/cgi-bin/submit.cgi?type=refine2.

MolProbity (n.d.). Protein structure validation. MolProbity. http://molprobity.biochem.duke.edu/.

SAVES v6.1 (n.d.). Structure validation server. UCLA. https://saves.mbi.ucla.edu/.

Expasy (n.d.). Swiss bioinformatics resource portal. SIB. https://www.expasy.org/.

ClusPro 2.0 (n.d.). Protein–protein docking server. ClusPro. https://cluspro.bu.edu/login.php.

PyMOL (n.d.). Molecular visualization. PyMOL. https://www.pymol.org/.