Journal of Underutilized Legumes
Vol. 7 (1): 12 - 21

Society for Underutilized Legumes
https://sulegumes.org/
e-ISSN: 2705-3776

Phylogenetic Analysis and Protein Structure Characterization of the *mat*K Gene in *Sphenostylis stenocarpa* and Related Legumes

[1]Edu N.E., [1,2]Edem U.L, [1]Osuagwu A.N., [3]Ojua E.O., [4]Aibuedefe O.E., [1]Oyohosuho P.J., [1]Emeagi L.I., and [1]Iheanetu C.N.

[1]Department of Genetics & Biotechnology, University of Calabar, P.M.B 1115, Calabar, Nigeria
[2]Animal Genomics and Bioresource Research Unit, Faculty of Science, Kasetsart University, Chatuchak, Bangkok 10900, Thailand.
[3]Department of Plant Science and Biotechnology, University of Nigeria, Nsukka, Nsukka Campus, P.M.B. 410001 Nsukka, Enugu State, Nigeria
[4]Centre for Forensic Science and DNA Studies, University of Benin, P.M.B. 1154, Edo State, Nigeria.

Email: edemlinus@unical.edu.ng

Received January 19, 2025
Accepted for publication January 22, 2025
Published January 23, 2025

Abstract

Legumes are essential for global nutrition and agriculture, providing significant protein, vital nutrients, and beneficial compounds. African yam bean (*Sphenostylis stenocarpa*), an important legume for agriculture and food security, faces challenges like high anti-nutritional factors, hard seed coats, long lifecycles, and photoperiod sensitivity. Their genetic diversity and that of related legumes remain underexplored. The Maturase K (*mat*K) gene, a chloroplast marker with a high substitution rate, is widely used in studying genetic diversity and species evolution. This study focuses on the *mat*K gene in legumes, specifically analysing *S. stenocarpa* and related species, to enhance understanding of their genetic diversity and potential for improvement. Nucleotide sequences for several leguminous species, including *S. stenocarpa, Sphenostylis angustifolia, Vigna aconitifolia, Vigna angularis, Vigna umbellata, Vigna mungo, Cajanus cajan, Phaseolus vulgaris, and Glycine max* were retrieved from NCBI database. Phylogenetic relationships were assessed using MEGA 6 software with Clustal W alignments and 1000 bootstrap resampling. The secondary and tertiary structures of proteins of the *mat*K gene were predicted using the GORIV and Phyre2 tools, respectively. Phylogenetic analysis revealed two primary clusters: one containing exclusively *P. vulgaris* with high bootstrap support, and another encompassing the remaining legumes, further divided into sub-clusters with *C. cajan* distinct from *Vigna* species. Structural analysis showed *S. stenocarpa* exhibited the highest percentage of alpha helix (36.54%), while *C. cajan* displayed the lowest alpha helix and highest random coil. Notably, *P. vulgaris* had the highest percentage of extended strands (35.21%). Tertiary structure predictions indicated that while *Vigna* species shared similar folding patterns, *P. vulgaris* and *C. cajan* had unique tertiary structures. These findings underscore significant evolutionary differences among the legumes and highlight the potential for genetic enhancement of these important crops.

Keywords: *mat*K Gene, Genetic diversity, Phylogenetic Analysis, Protein Structure, Leguminous Plants

Introduction

Legumes represent a diverse plant family that offers a myriad of resources essential for human consumption and agriculture. These plants are significant sources of protein, oils, minerals, and various nutraceuticals [1,2,3]. Grain legumes, including African yam bean (*S. stenocarpa*) and adzuki bean (*V. angularis*), serve as essential staple foods, providing more than 33% of the dietary protein consumed by humans. Key contributors to this protein supply include species like cowpea (*V. unguiculata*), pigeon pea (*C. cajan*), and common bean (*P. vulgaris*). [4,5,6]. Additionally, refined oils, particularly from soybeans (*G. max*), have widespread industrial uses, including in paints, diesel fuels, and solvents. Legumes also accumulate significant phytochemicals like isoflavonoids, linked to health benefits [7,8]. A notable characteristic of legumes is their ability to form symbiotic relationships with soil microbes, particularly rhizobia, which facilitate nitrogen fixation through the development of root nodules [9,10,11]. This capability is crucial for enhancing soil fertility and increasing agricultural productivity, especially in sub-Saharan Africa, where over 60% of the population relies on these crops for protein intake [11,12].

Regrettably, the agro-biodiversity of many leguminous species has declined significantly in Nigeria, largely due to a lack of awareness about their economic potential, inefficient propagation, inadequate processing methods, and limited market access [3,13]. Furthermore, minimal genetic improvement has been aimed at enhancing these crops' agronomic and nutritional quality. Understanding these legumes' genetic makeup and protein profiles is essential for promoting their improvement and utilisation [14,15]. The characterization of existing germplasm can be effectively achieved through robust genetic diversity analyses, employing molecular and sequence data. [3,14,16,17,18,19,20].

Recent studies have highlighted the utility of chloroplast, mitochondrial, and nuclear genes for elucidating evolutionary trends at the genus level [21,22,23,24,25]. By using a combination of chloroplast, mitochondrial, and nuclear genes, scientists can create more precise phylogenetic trees. This approach enhances the understanding of evolutionary connections, helps clear up taxonomic uncertainties, and enables the differentiation of species that are closely related within the same genus. [24,25] Among these, the maturase K (*mat*K) gene has emerged as a significant marker for plant molecular systematics and evolution due to its rapid evolutionary rates [26,27,28,29] Rapid nucleotide substitutions and limited occurrences of frameshift insertions and premature stop codons necessitate careful analysis of *mat*K's functionality in various plant species [30]. The RNA transcripts of several chloroplast genes, including *trn*K and *trn*A, depend on *mat*K for proper functioning [29,31,32,33]. Given the significance of *mat*K in phylogenetic studies, this research aims to characterize the *mat*K gene and analyze protein structural variations in selected legumes, specifically *S. stenocarpa, S. angustifolia, V. aconitifolia, V. angularis, V. umbellata, V. mungo, C. cajan, P. vulgaris, and G. max.* By employing bioinformatics tools, this study seeks to assess how the secondary and tertiary structures of *mat*K gene vary among these species, providing insights into their genetic relationships and potential for improvement.

Materials and Methods

Study location

The in-silico analysis was conducted in the Bioinformatics Laboratory of the Department of Genetics and Biotechnology at the University of Calabar, Calabar.

Retrieval of Nucleotide and Amino Acid Sequences

Nucleotide sequences for African yam beans and related legumes were downloaded from the National Center for Biotechnology Information (NCBI) database in FASTA format. The analysis

considered sequences from species including *S. stenocarpa;* AY582977.1*, S. angustifolia;* JN008190.1, AY582978.1*, V. aconitifolia;* MH311596.1, MH311594.1, MH311592.1*, V. angularis* MH311581.1, MH311577.1, MH311580.1*, V. umbellate* MH311601.1, MH311600.1, MH311598.1 *, V. mungo;* MH311602.1, MH311606.1, MH311605.1*, C. cajan* JN228940.1, OQ289261.1, MN166627.1 *, P. vulgaris;* LC578842.1, LC578844.1, LC578843.1*,* and *G. max;* MW316063.1, MT23931.7.1, MH659986.1 .

Phylogenetic Analysis of matK gene in African yam bean and related legumes.
Molecular Evolutionary Genetic Analysis (MEGA) software version 6.0 was used to analyse the phylogenetic relationships. Clustal W was utilised for nucleotide sequence alignment, with gaps excluded from the analysis. Phylogenetic trees were constructed using 1000 bootstrap replicates [34].

Prediction of Secondary and Tertiary Protein Structures
The GORIV online tool used amino acid sequences of the *mat*K gene to predict secondary structures. The canonical amino acid sequences obtained from the NCBI database through the Phyre2 server were used to model tertiary protein structures as described by Edem et al. [34] The protein chain patterns in *mat*K gene from the legume varieties were analysed using the Simple Modular Architecture Research Tool, available at https://smart.embl-heidelberg.de/. To predict the three-dimensional structure of the proteins, Phyre2 online software was employed, utilising the reference protein sequence obtained from the NCBI database. The samples chosen for analysis of protein patterns and structures were selected based on their classification in the phylogenetic tree.

Results

Phylogenetic of selected legumes
Figure 1 illustrates the phylogenetic relationships among the selected legume species. There were two major clusters. The first cluster was made of all the samples of *P. vulgaris* with 100% bootstrap. The second major cluster consisted of all the remaining legumes. In the second major cluster, two sub-clusters were generated where *C. cajan* occupied one of the Vigna genes found on some sub-clusters. In general, the legumes were grouped into specific clusters based on their relatedness at both the generic and specific levels.

Protein structure variation
Table 1 and Figures 2-10 display the findings on protein structure variations at both the secondary and tertiary levels. From Table 1, there were variations in the secondary protein structure as revealed by the percentage of alpha helix, extended strand, and random coil. All the Vigna legumes had similar percentages of alpha helix, extended strand, and random coils. *S. stenocarpa* had the highest percentage of alpha helix (36.54%) while *C. cajan* had the lowest. However, *C. cajan* had the highest random coil. The extended strand was more in the amino acid square of *P. vulgaris* (35.21%). Similar results were obtained for the tertiary protein structure of the selected legumes. All species from the *Vigna* genus had similar folding showing the alpha helix, extended strand as random coil. From Figure 7, it was obvious that *P. valgarius* had quite dissimilar folding patterns from the other species of legumes. Similarly, *C. cajan* also shows a different folding pattern from other legumes used in the present study.

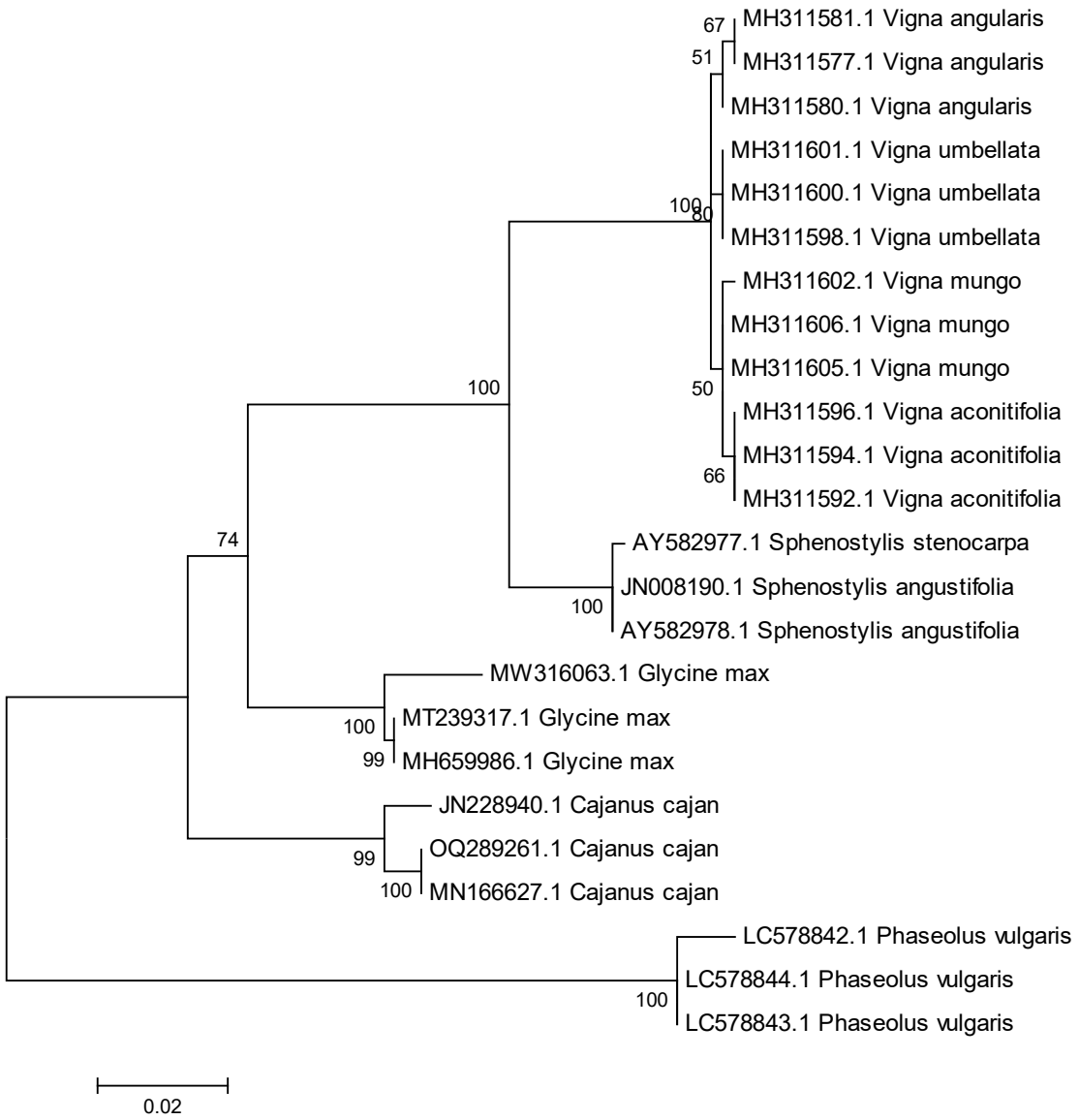

Figure 1: Phylogenetic relationship between selected legumes based on *mat*K gene

Table 1: Subunits of the secondary protein structure in the *matK* gene of African yam bean and its related legume species.

Subunit (%)	*S. stenocarpa*	*S. angustifolia*	*P. vulgaris*	*V. angularis*	*V. mungo*	*V. aconitifolia*	*V. umbellata*	*C. cajan*	*G. max*
Alpha helix	36.54	35.58	25.29	29.73	29.61	29.13	29.13	17.70	24.76
Extended strand	27.40	29.33	36.21	27.80	27.67	30.58	27.67	31.10	35.24
Random coil	36.06	35.10	38.51	42.44	42.44	40.29	43.20	51.20	40.00

Figure 2: The tertiary structure of the *mat*K protein in *V. angularis*.

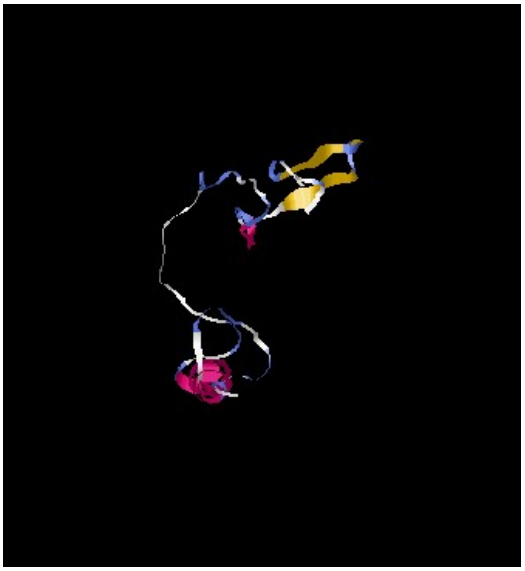

Figure 3; The tertiary structure of the *mat*K protein in *V. mungo*

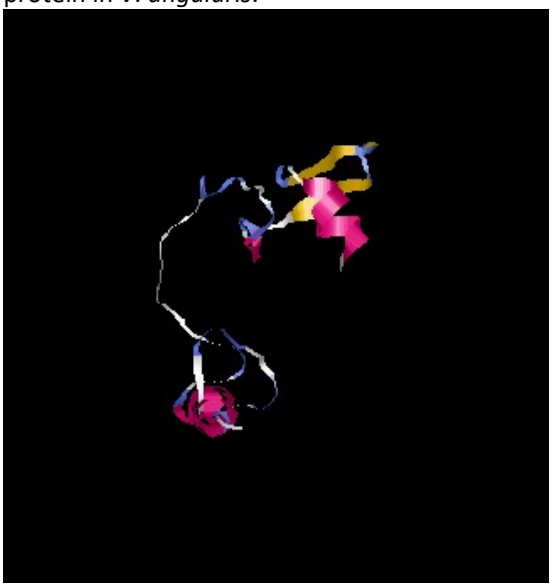

Figure 4: The tertiary structure of the *mat*K protein in *V. aconitifolia*.

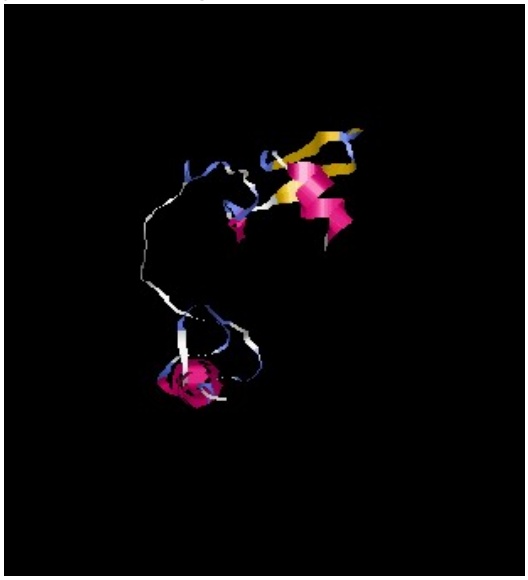

Figure 5: The tertiary structure of the *mat*K protein in *V. umbellata*

Figure 6: The tertiary structure of the *mat*K protein in *G. max*

Figure 7: The tertiary structure of the *mat*K protein in *P. vulgaris*

Figure 8: The tertiary structure of the *mat*K protein in *C. cajan*

Figure 9: The tertiary structure of the *mat*K protein in *S. stenocarpa*

Figure 10: The tertiary structure of the *mat*K protein in *S. angustifolia*

Discussion

The phylogenetic analysis presented in Figure 1 delineates two major clusters among the selected legumes. The first cluster is exclusively composed of *P. vulgaris* with 100% bootstrap support, indicating a strong and reliable monophyletic grouping for this species. This finding is consistent with Edem and Osuagwu [15], who observed that *P. vulgaris* consistently forms a distinct clade with *rbc*L in legume phylogenies. The second major cluster includes the remaining legumes, which further segregates into two sub-clusters. One sub-cluster contains *C. cajan*, while the other encompasses various species from the *Vigna* genus. The clear distinction of *C. cajan* into its sub-cluster highlights its unique evolutionary lineage, this clustering reinforces our understanding of the evolutionary relationships within the legume family.

The variations in secondary structure composition among the studied legume species offer insights into their structural adaptations and potential functional roles [35]. The high alpha-helix content in *S. stenocarpa* (36.54%) suggests a structural adaptation that prioritizes stability and compactness. Alpha helices are often associated with robust protein stability and a defined structural framework [36], which may be critical for the physiological functions and resilience of *S. stenocarpa* in its ecological niche.

In contrast, *C. cajan* exhibited the lowest alpha-helix percentage (17.70%) and the highest random coil percentage (51.20%), indicative of a more flexible protein structure. Random coils are unstructured regions that can allow greater conformational freedom and adaptability [37]. This flexibility might enable *C. cajan* to respond to environmental stresses, such as drought, or play a role in its metabolism, which may demand dynamic interactions with other biomolecules. *P. vulgaris*, with the highest percentage of extended strands (35.21%), highlights another form of structural specialization. Extended strands, commonly forming beta-sheets, are known for their role in creating stable, planar protein structures that may be involved in molecular recognition or binding [38]. This feature could align with specific functional demands of *P. vulgaris*, such as its enzymatic activities or interaction with other biomolecules essential for its growth and reproduction.

The observed tertiary structures of the legume species exhibit relatively consistent folding patterns, suggesting a high degree of structural conservation among these legumes. This consistency could reflect conserved evolutionary pressures or functional requirements associated with their shared ecological niches and physiological roles [39]. In contrast, *P. vulgaris* displays a notably different tertiary folding pattern, as highlighted in Figure 6. This deviation may indicate unique adaptive strategies or evolutionary divergence, potentially driven by specific environmental factors, distinct metabolic demands, or differences in selective pressures that have influenced its genome and protein folding mechanisms [40]

Similarly, the unique tertiary structures observed in *C. cajan* and *S. stenocarpa* (Figures 7 and 8) underscore their distinct evolutionary trajectories. These structural distinctions could be attributed to gene expression, sequence variation, or environmental interaction differences [41]. For instance, *C. cajan* (pigeon pea) is well-adapted to drought-prone areas, which may influence its molecular adaptations [42] *S. stenocarpa*, being less studied, represents a valuable model for exploring underutilized crop diversity, as its structural uniqueness might relate to untapped genetic traits for resilience or nutritional value [43].

In conclusion, this research underscores the importance of understanding genetic and protein structural diversity in legumes, which could enhance the development of improved varieties with better resilience, nutritional quality, and adaptability, ultimately contributing to food security, particularly in sub-Saharan Africa. Further research should examine the molecular mechanisms

behind the structural adaptations in *S. stenocarpa* and *C. cajan*, focusing on genetic markers and regulatory pathways. Functional genomics techniques can provide insights into their ecological and nutritional roles. Additionally, studying structural variations in legumes like *P. vulgaris* could aid breeding efforts to improve crop yield, stress resistance, and nutritional quality.

Authors' contribution(s)
Ndem Enyogor Edu supervised the experiment
Uduak Linus Edem designed the experiment
Aniefiok Ndubuisi Osuagwu performed the statistical analysis
Eugene Obashi Ojua, sort for computational packages
Osagie Eremwanarue Aibuedefe reviewed the manuscript
Phillip, Julius Oyohosuho performed the methodology
Lasbrey Ikenna Emeagi carried out the literature search
Cynthia Nkeiruka Iheanetu, reviewed the statistical analysis

Acknowledgements
We extend our gratitude to all those who contributed their efforts to this research work.

Declaration of Competing Interest
The authors declare no conflicts of interest.

Funding Information
There was no funding provided for this research

References

1. Guo F, Danielski R, Santhiravel S, Shahidi F. Unlocking the nutraceutical potential of legumes and their by-products: Paving the way for the circular economy in the agri-food industry. *Antioxidants*. 2024; 13:636. Available from: https://doi.org/10.3390/antiox13060636
2. Lisciani S, Marconi S, Le Donne C, Camilli E, Aguzzi A, Gabrielli P, et al. Legumes and common beans in sustainable diets: Nutritional quality, environmental benefits, spread and use in food preparations. *Front in Nutrition*. 2024; 11:1385232. Available from: https://doi.org/10.3389/fnut.2024.1385232
3. Edem UL, Osuagwu AN, Edu NE, Emeagi LI, Ekerette EE, Ogbeche AA, et al. Assessment of genetic diversity analysis in African yam bean (*Sphenostylis stenocarpa Hochst Ex. A. Rich Harms*) using qualitative and quantitative attributes. *J of Adv in Bio & Biotech*. 2024;27(10):122–31. https://doi.org/10.9734/jabb/2024/v27i103535
4. Ganiyu O. Antioxidant properties of some commonly consumed and underutilized tropical legumes. *Eur Food Res Technol*. 2006;224(1):61–5. Available from: https://doi.org/10.1007/s00217-006-0289-x
5. James S, Nwabueze TU, Onwuka GI, Ndife J, Usman MA. Chemical and nutritional composition of some selected lesser-known legumes indigenous to Nigeria. *Heliyon*. 2020;6(11): e05497. Available from: https://doi.org/10.1016/j.heliyon.2020.e05497
6. Fukah FK, Magubika AJ, Tryphone GM, Nassary EK. Meta-analysis of legumes and groundnut production trends and variability in the Global South. *J Agric Food Res*. 2024; 18:101501. Available from: https://doi.org/10.1016/j.jafr.2024.101501
7. Kinney AJ, Clemente TE. Modifying soybean oil for enhanced performance in biodiesel blends. *Fuel Process Technol*. 2005;86(10):1137–47. Available from: https://doi.org/10.1016/j.fuproc.2004.11.008
8. Le, X. T., Lan Vi, V. L., Toan, T. Q., Bach, L. G., Truc, T. T., & Hai Ha, P. T. Extraction Process of Polyphenols from Soybean (*Glycine max* L.) Sprouts: Optimization and Evaluation of Antioxidant Activity. *Processes*, 2019, *7*, 489. https://doi.org/10.3390/pr7080489

9. Wang Q, Liu J, Zhu H. Genetic and molecular mechanisms underlying symbiotic specificity in legume-rhizobium interactions. *Front Plant Sci*. 2018; 9:313. doi:10.3389/fpls.2018.00313
10. Coba de la Peña T, Fedorova E, Pueyo JJ, Lucas MM. The symbiosome: Legume and rhizobia co-evolution toward a nitrogen-fixing organelle? *Front Plant Sci*. 2018; 8:2229. doi:10.3389/fpls.2017.02229
11. Fahde S, Boughribil S, Sijilmassi B, Amri A. Rhizobia: A promising source of plant growth-promoting molecules and their non-legume interactions: Examining applications and mechanisms. *Agriculture*. 2023; 13:1279. doi:10.3390/agriculture13071279
12. Ladha JK, Peoples MB, Reddy PM, Biswas JC, Bennett A, Jat ML, Krupnik TJ. Biological nitrogen fixation and prospects for ecological intensification in cereal-based cropping systems. *Field Crops Res*. 2022; 283:108541. doi: 10.1016/j.fcr.2022.108541
13. Khan MMH, Rafii MY, Ramlee SI, Jusoh M, Al-Mamun M. Bambara groundnut (*Vigna subterranea* L. Verdc): A crop for the new millennium, its genetic diversity, and improvements to mitigate future food and nutritional challenges. *Sustainability*. 2021; 13:5530. doi:10.3390/su13105530
14. Aziz MA, Masmoudi K. Molecular breakthroughs in modern plant breeding techniques. *Hortic Plant J*. 2024. doi: 10.1016/j.hpj.2024.01.004
15. Edem UL, Osuagwu AN. Evaluation of secondary and tertiary protein structure variation in *rbc*L gene of selected legumes using a computational approach. *J Proteomics Bioinform*. 2023; 16:642.
16. Kilian B, Graner A. NGS technologies for analyzing germplasm diversity in genebanks. *Brief Funct Genomics*. 2012;11(1):38–50. doi:10.1093/bfgp/elr046
17. Edem UL, Osuagwu AN, Udensi OU. Assessment of genetic diversity in *Sphenostylis stenocarpa* (Hochst Ex. A Rich Harms) using morphological markers. *Asian J Biol*. 2023;19(4):74-85. doi:10.9734/ajob/2023/v19i4378
18. Emeagi LI, Edem UL, Osuagwu AN, Ker SS. Cluster and principal component analysis of African yam bean (*Sphenostylis stenocarpa* Hochst. ex A. Rich. Harms) using morphological attributes. *Int J Sci Acad Res*. 2024;5(5):7518-7526.
19. Osuagwu AN, Edem UL. Relationship between salt (NaCl) stress and yield components in three varieties of African yam bean (*Sphenostylis stenocarpa*). *Asian J Biol*. 2021;13(1):30-176. doi:10.9734/ajob/2021/v13i130176
20. Osuagwu A, Edem U, Edu N, Phillip J, Emeagi L, Iheanetu C, et al. Unveiling the genetic diversity of aerial yam (*Dioscorea bulbifera* L.): Morphological insights for sustainable crop improvement. *Annu Res Rev Biol*. 2024;39(12):78-89. doi:10.9734/arrb/2024/v39i122173
21. Keller J, Rousseau-Gueutin M, Martin GE, Morice J, Boutte J, Coissac E, et al. The evolutionary fate of the chloroplast and nuclear *rps16* genes as revealed through the sequencing and comparative analyses of four novel legume chloroplast genomes from *Lupinus*. *DNA Res*. 2017;24(4):343-58. doi:10.1093/dnares/dsx006
22. Distefano G, La Malfa S, Currò S, Las Casas G, Wünsch A, Gentile A. HRM analysis of chloroplast and mitochondrial DNA revealed additional genetic variability in *Prunus*. *Sci Hortic*. 2015; 197:124-29. doi: 10.1016/j.scienta.2015.09.020
23. Zhou J, Zhang S, Wang J, et al. Chloroplast genomes in *Populus* (*Salicaceae*): Comparisons from an intensively sampled genus reveal dynamic patterns of evolution. *Sci Rep*. 2021; 11:9471. doi:10.1038/s41598-021-88160-4
24. Maughan PJ, Chaney L, Lightfoot DJ, et al. Mitochondrial and chloroplast genomes provide insights into the evolutionary origins of quinoa (*Chenopodium quinoa* Willd.). *Sci Rep*. 2019; 9:185. doi:10.1038/s41598-018-36693-6
25. Rousseau-Gueutin M, Keller J, Ferreira de Carvalho J, Aïnouche A, Martin G. The intertwined chloroplast and nuclear genome coevolution in plants. *IntechOpen*. 2018. doi:10.5772/intechopen.75673

26. Barthet MM, Hilu KW. Expression of *matK*: Functional and evolutionary implications. *Am J Bot*. 2007;94(8):1402-12. doi:10.3732/ajb.94.8.1402
27. Selvaraj D, Sarma RK, Ramalingam S. Phylogenetic analysis of chloroplast *mat*K gene from Zingiberaceae for plant DNA barcoding. *Bioinformation*. 2008;3(1):24-27. doi:10.6026/97320630003024
28. Udensi OU, Ita EE, Ikpeme EV, Ubi G, Emeagi LI. Sequence analysis of maturase K (*mat*K): A chloroplast-encoding gene in some selected pulses. *Glob J Pure Appl Sci*. 2017; 23:213-30. doi:10.4314/gjpas. v23i2.2
29. Letsiou S, Madesis P, Vasdekis E, Montemurro C, Grigoriou ME, Skavdis G, et al. DNA barcoding as a plant identification method. *Appl Sci*. 2024; 14:1415. doi:10.3390/app14041415
30. Mower JP, Ma P-F, Grewe F, Taylor A, Michael TP, VanBuren R, et al. Lycophyte plastid genomics: Extreme variation in GC, gene and intron content and multiple inversions between a direct and inverted orientation of the rRNA repeat. *New Phytol*. 2019;222(2):1061-75. doi:10.1111/nph.1565
31. Muino JM, Ruwe H, Qu Y, Maschmann S, Chen W, Zoschke R, et al. *matK* impacts differential chloroplast translation by limiting spliced tRNA-K(UUU) abundance. *Plant J*. 2024. doi:10.1111/tpj.16945
32. del Campo EM. Post-transcriptional control of chloroplast gene expression. Gene Regul Syst Bio. 2009; 3:31–47. doi:10.4137/grsb.s2080
33. Zhang Y, Tian L, Lu C. Chloroplast gene expression: Recent advances and perspectives. *Plant Commun.* 2023;4(5):100611. doi: 10.1016/j.xplc.2023.100611
34. Edem UL, Osuagwu AN, Edu NE, et al. Assessing genetic diversity in squash pumpkin (*Cucurbita moschata*) through computational analysis of plastid genes. *Trop J Nat Prod Res.* 2024;8(11):9310–4. doi:10.26538/tjnpr/v8i11.46
35. Carbonaro M, Maselli P, Nucara A. Structural aspects of legume proteins and nutraceutical properties. *Food Res Int.* 2014; 76:19–30. doi: 10.1016/j.foodres.2014.11.007
36. Swanson CJ, Sivaramakrishnan S. Harnessing the unique structural properties of isolated α-helices. J Biol Chem. 2014;289(37):25460–7. doi:10.1074/jbc. R114.583906
37. Berlow RB, Dyson HJ, Wright PE. Expanding the paradigm: Intrinsically disordered proteins and allosteric regulation. *J Mol Biol.* 2018;430(16):2309–20. doi: 10.1016/j.jmb.2018.04.003
38. Eswar N, Ramakrishnan C, Srinivasan N. Stranded in isolation: Structural role of isolated extended strands in proteins. Protein Eng. 2003;16(5):331–9. doi:10.1093/protein/gzg046
39. Udensi UO, Emeagi IL, Daniel VE, et al. Insight into the genetic diversity of pigeon pea (*Cajanus cajan L. Millsp.*) and cowpea (*Vigna unguiculata L. Walp.*) germplasm cultivated in Nigeria based on rbcl gene region. *Genet Resour Crop Evol.* 2022; 69:2231–48. doi:10.1007/s10722-022-01373-y
40. Lee SC, Gepts PL, Whitaker JR. Protein structures of common bean (*Phaseolus vulgaris*) alpha-amylase inhibitors. *J Agric Food Chem.* 2002;50(22):6618–27. doi:10.1021/jf020189t
41. Smartt J, Haq N, editors. New crops and uses: Their role in a rapidly changing world. Blackwell Publishing; 2007.
42. Musokwa M, Mafongoya P. Pigeonpea yield and water use efficiency: A savior under climate change-induced water stress. *Agronomy*. 2021;11(5):5. doi:10.3390/agronomy11010005
43. Shitta NS, Unachukwu N, Edemodu AC, et al. Genetic diversity and population structure of an African yam bean (*Sphenostylis stenocarpa*) collection from IITA GenBank. *Sci Rep.* 2022;12(1):4437. doi:10.1038/s41598-022-08271-4