Comprehensive Proteomics Analysis of Function Study of AKR4C14, an Aldo-keto Reductase from Thai Jasmine Rice by Applying NGS with Biopython

Shephali Sachan[1], Uma Kumari[1], Murugesan AnbuMegala[2]
[1]Project Trainee, Bioinformatics Project and Research Institute, Noida-201301,India
[1]Senior Bioinformatics Scientist, Bioinformatics Project and Research Insitute, Noida - 201301, India
[2]Senior Lecturer,St.Eugene University,Zambia
Corresponding Author: Uma Kumari(uma27910@gmail.com)

Abstract

Aldo-keto reductases (AKRs) are important enzymes involved in various metabolic pathways, including the detoxification of aldehydes and ketones.AKR4C14, which is an aldo-keto reductase found in Thai Jasmine rice (Oryza sativa), has been linked to crucial physiological processes. However, its functional roles have not been extensively studied. In this research, we will use next-generation sequencing (NGS) technologies along with biopython tools to conduct a thorough proteomics analysis of AKR4C14. Aldo-keto reductases (AKRs) are oxidoreductase enzymes that metabolize aldehydes and ketones into their corresponding alcohols in a NADPH/$NADP^+$ dependent manner. Without the cofactor, the apo structure of indica rice AKR4C14 was likewise determined, displaying the stabilized open conformation. Our goal is to understand the functional roles of AKR4C14 and how it interacts with metabolic pathways in rice by integrating proteomic data.

Keywords: Oryza sativa, AKR4C14, Next-Generation Sequencing, Biopython, Proteomics

Introduction

The continuous, progressive development in technological innovations has greatly benefited the biotechnology and molecular biology sector, especially in the realms of high-throughput sequencing and computational analysis. These advancements have made it possible to regulate various superior quality traits and simultaneously detect and remove harmful or undesirable traits in crops. Ultimately, this leads to a path towards strong, sustainable agriculture (Fang et al., 2016).In light of current circumstances such as the increasing global population (approximately 8.2 billion), declining crop production (by 38%), rising levels of malnutrition (around 9%), and a hunger crisis affecting approximately 850 million people, it is essential to prioritize sustainable agriculture in order to ensure food security for the future (Ashraf et al., 2022).

The innovation of Next-Generation Sequencing (NGS) has greatly contributed to crop improvement by revealing the genetic potential of wild and domesticated varieties. NGS has revolutionized genomic and transcriptomic research by enabling high-throughput, detailed, and accurate sequencing of DNA, RNA, and protein. It facilitates the analysis of gene expression profiles, detection of genetic variationand epigenetic modifications, and identifying and quantifying transcripts associated with specific proteins. This approach allows researchers to identify the differential expression of genes under various conditions, including stress, and link changes in gene expression to protein function and regulation, ultimately leading to the generation of specific crop phenotypes(Ashraf et al., 2022).

The integration of powerful bioinformatics tools such as Biopython with NGS allows researchers to efficiently handle large datasets, perform complex analyses, and interpret results with greater precision. Biopython is an open-source toolkit that offers a range of tools for processing biological data, including sequence alignment, data visualization, and functional analysis. Its capabilities are particularly useful for tasks such as parsing and aligning sequence data, analyzing expression levels,

and integrating transcriptomic and proteomic information. The toolkit facilitates the exploration of functional relationships between genes and proteins, enabling a more comprehensive understanding of their roles in biological processes (Mu et al., 2022).

Aldo-keto reductases (AKR4C14) enzyme found in Thai Jasmine rice (Oryza sativa) by using different NGS technique and Biopython tool. This enzyme is known for its role in detoxifying reactive aldehydes and ketones. In plants, it plays crucial roles in stress responses, hormone regulation, and secondary metabolism. The genetic potential of AKR4C14 has not been fully explored. Therefore, this study aims to provide deeper insights into the enzyme's activities and interactions to better understand its importance in rice physiology and stress tolerance mechanisms (Songsiriritthigul et al., 2020).

Material and Methods

1. Sample Retrieval

The 3D structure of experimental protein aldo-keto reductase enzyme, AKR4C14 of Thai Jasmine Rice or Thai Hom Mali Rice (Oryza sativa L. ssp. Indica cv. KDML105) was extracted from Protein Database Bank (PDB) under the PDB ID 6KBL. All required related fundamental information was also studied and extracted from PDB (Table 1).

2. BLAST p Search

The extracted AKR4C14 protein as 6KBL was then subjected to blast p for obtaining and identifying the maximum aligned sequences. The FASTA sequence of 6KBL protein was retrieved using MMDB (Molecular Modelling Database) tool of NCBI (National Center of Biotechnology Information). This sequence was inserted in the interface of blast p tool. Rest all the parameters in blast p search were remained as default settings.

Table 1: Fundamental Information of AKR4C14 Protein

Classification	OXIDOREDUCTASE
Organism(s)	Oryza sativa Indica Group
Expression System	Escherichia coli BL21(DE3)
Mutation(s)	No
Method	X-RAY DIFFRACTION
Resolution	1.70 Å
R-Value Free	0.194
R-Value Work	0.166
R-Value Observed	0.167
Total Structure Weight	35.01 kDa
Atom Count	2,674
Modelled Residue Count	314
Deposited Residue Count	314
Unique protein chains	1

3. Multiple Sequence Alignment (MSA)

MSA was performed on the basis of blast p results using CLUSTAL OMEGA software. The selected sequences along with query (AKR4C14) sequence runin the CLUSTAL OMEGA interface. The alignment can beidentifiedby identical sign (*).Colour coding in sequence narrate the brief quantitative and qualitative information about the conserved region among these sequences.

4. Next Generation Sequencing

Protein Network and Enrichment Analysis

The STRING software appliedto identify and study all possible protein-protein interactions of ourtarget protein sequence,AKR4C14.The establishment of interaction is based on the available data, including experimental evidence, computational predictions, and known interactions from curated databases. The FASTA sequence of 6KBL is uploaded in the STRING interface and proceeded further to identify the respective interacting protein and organism.The protein-protein interaction networks were constructed to explore potential interactions with other metabolic enzymes.Each interaction was assigned with confidence scores (0.15 – 0.9) from lowest to highest score, reflecting the reliability of each predicted interaction.Higher score (≥ 0.4, medium confidence score) indicatethe stronger evidenceof the precision of results. This information will further provide the backbone for gene ontology.

Functional Analysis

The sequence AKR4C14 was submitted to Interproscan software to annotate the sequence structurally and functionally. Interproscan is the next generation sequencing (NGS) tool which scans any protein or nucleic acid sequence for annotation.Gene Ontology (GO) and pathway analysis were performed to determine the biological functions and pathways associated with AKR4C14.

Additionally, SCANPROSITE software applied to analyse the target protein sequenceAKR4C14for functional motifs and domains. This tool matches the input sequences against the PROSITE database, which contains curated patterns associated with specific biological functions. The sequence was submitted to ScanProsite, and focused on identifying the significant hits with zero confidence level (Higher reliability). This could suggest potential roles or interactions of the proteins under study.

5. Biopython

Biopython is a powerful toolkit for bioinformatics, was used to process NGS data. Several biopython tasks were performed through Google Colab open access application. Following steps included:

1. Parsing FASTA files and aligning reads to the rice genome.
2. Proteomic data to assess the functional significance of AKR4C14.

Results and Discussions

1. Similarity Analysis of Thai Jasmine Rice Proteomic Sample

The amino acid sequence of 6KBL (AKR4C14)protein was extracted using MMDB tool is given in Figure 1.

>pdb|6KBL|A Chain A, Aldo-keto reductase
AAPMAKHLVLNTGAKIPSVGLGTWQSDPGVVGNAVYAAVKAGYRHIDCASAYNNEKEVGLALK
KLFEEGV
VKRADLFITSKLWCDHHAPEDVPEALDATLNDLQLEYLDLYLIHWPFRTKKGSSIGKPESYLPPDIP
STW
AAMEKLYDSGKSRAIGVSNFSSKKLGDLLAAARVPPAVDQVECHPGWQQMKLHNFCQSTGIHLS
AYSPLG
SPGSTFMNGNVLKEPIIISIAEKLGKTPAQVALRWNIQMGHSVLPKSVSEERIKQNLDVYDWSIPED
LLA
KFSEIKQVRLLRGNFIVNPQSVYKTHEELWDGEI

Figure 1: Amino Acid Sequence of 6KBL Protein

The blast p results of 6KBL protein displayed several amino acid sequences related to cereal family with appropriate query cover, E-value and percentage identity. Out of which two proteins related to other species viz: 5JH2_A (Zea mays) and 7W1X_A (Echinochloa colona) were selected. These sequences were found not putative, experimentally derived, and closest or highly aligned with the query sequence 6KBL showing highest query cover and percentage identity, and lowest (zero) E-value (Table 2).

Table 2: Query Cover, E-value and Percentage Identity of Sequences with 6KBL

Species	NCBI ID, PDB ID	Query Cover (%)	E-value	Percentage Identity (%)	Description
Zea mays	5JH2_A, 5JH2	99	0.0	86.50	AKR4C7 (Aldose reductase)
Echinochloa colona	7W1X_A, 7W1X	98	0.0	87.38	AKR4-1 (Oxidoreductase)

(Sachan et al., 2024)

The blast program appeared to be most powerful tool used to find out the similar or homologous sequences from database for the target sequence. These homology regions are always expressed in percentage. The sequences with the highest query cover (nearer to 100%) and percentage identity (<70%), and lowest E-value (nearer to zero) shows maximum alignment with the target sequence(Samal et al., 2021). In the present study, the selected 5JH2 and 7W1X sequences of Zea mays and Echinochloa colona respectively displays maximum alignment with the target sequence 6KBL (Table 1) according to obtained query cover, E-value (Expected value) and percentage identity scores. This tool is widely used worldwide because of flexible search, reliable, accurate and fast speed results, and quality statistical report (Samal et al., 2021).

2. Multiple Sequence Alignment

The Multiple Sequence Alignment (MSA) performed among 6KBL, 5JH2 and 7W1X using CLUSTAL OMEGA software. The results presented the maximum alignment depicting about the conserved sequence sites related to unique and common function (Figure 2). The maximum conserve site cover by blue colour followed by green, magenta, yellow and red colour respectively. Blue colour depicts the richness of hydrophobic amino acid residues (A, I, L,M, F, W,V), green colour for polar (N,Q,S,T),magenta colour for negative charge (E,D), yellow for prolines and red colour displays the abundance of positive charge (K,R) amino acid residues. Furthermore, the scores displaying in the guide tree and phylogenetic tree clearly indicates that the 5JH2 and 7W1X sequences are more closely related to each other and also shares close evolutionary relationship.

The CLUSTAL OMEGA analysis highlights the presence of conserved sites that are critical for their unique and common functions, essential for understanding their biological roles (Sievers & Higgins, 2018).This analysis predicts the protein is likely to be a **globular protein** with a well-defined structure, featuring:

- A **hydrophobic core** formed by the blue (hydrophobic) residues.
- **Surface regions** enriched with polar and charged residues that interact with the solvent and other molecules.
- Potential **active or binding sites** formed by polar and charged residues, possibly in conjunction with the conserved hydrophobic core.

The guide tree and phylogenetic tree results provide further insights into the evolutionary relationships among these proteins(Sievers & Higgins, 2018). The close relationship between the 5JH2 and 7W1X sequences suggests they may have diverged from a common ancestor more recently than 6KBL. This close evolutionary relationship may reflect similar functional roles or mechanisms, potentially indicating a shared pathway or biological process. In contrast, the greater distance of 6KBL from the other two sequences suggests that it may have evolved different characteristics or functions.CLUSTAL OMEGA is the basic tool to obtain the information of align regions and evolutionary relationship among the DNA or protein sequences(Sievers & Higgins, 2018). To reflect the consensus and conserved residues found in the T4HNR protein (tetrahydroxynaphthalene reductase) members, complexed with NADP(H) and pyroquilon) fungicide employed against tomato vascular wilt (Fusarium oxysporumf.sp. lycopersici (FOL)), multiple sequence alignment was done(Aamir et al., 2018).

Similarly, conserved regions were identified and phylogenetic tree was constructed for eight Arabidopsis FRO (ferric reduction oxidase) proteins involved in homeostasis, tolerance and signalling networks responses against several abiotic stresses(Aamir et al., 2018).Fourteen sequences of L. POR (NADPH:protochlorophyllide (Pchlide) oxidoreductase, a key enzyme of chlorophyll biosynthesis) from Pisum sativum (Sameer et al., 2021) and eleven sequences of OsARD4 (acireductone dioxygenase, promotes secondary roots development) from eleven different species (Ramanathan et al., 2018)were also analysed for similarities with this software.

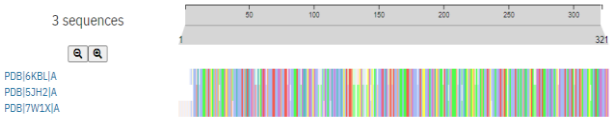

Figure 2: Multiple Sequence Alignment (MSA) Analysis of AKR4C14

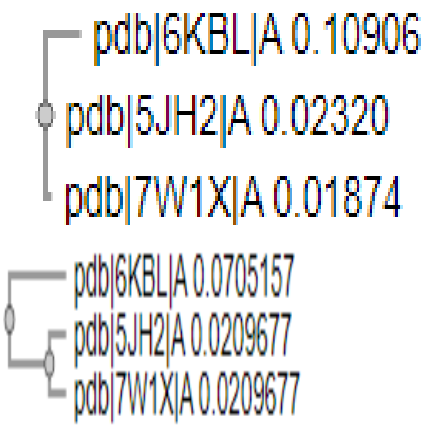

Figure 3: Guide Tree **Figure 4:** Phylogenetic Tree

3. Protein Network and Enrichment Analysis of AKR4C14

The string analysis of AKR4C14 protein showed several networks of interacting proteins out of which one protein PDC2, Pyruvate decarboxylase 2 (belongs to the TPP enzyme family) showed putative homology according to curated databases by 0.913 confidence level.This signifies the strong confidence in the accuracy of our results, indicating that the estimation method is generally reliable.Therefore, PDC2, the only characterized protein along with various other uncharacterized protein showed highest interaction with the AKR4C14 protein. The highest interaction (lowest confidence level) was observed by0.929 confidence level with an uncharacterized protein (Figure 5&6).

The interactome profiling through string database displayed the homology outcome with Pyruvate decarboxylase 2 (PDC2) enzyme. The network or pathways of both the enzymes were analysed connected. Bothinvolved in metabolic pathways processing carbon compounds but with the distinct roles. PDC2 contributes in the fermentation pathway and helps regenerate NAD^+ from NADH, which is essential for maintaining the glycolytic pathway under anaerobic conditions. While aldo-keto reductase enzyme reduces aldehydes and ketones to their corresponding alcohols, playing a crucial role in detoxification and metabolism of sugars, steroids, and lipids. They are interconnected through the broader metabolic pathways where pyruvate (produced from glycolysis) can enter various pathways, including those involving aldehyde or ketone intermediates that AKR might act upon. The confidence scores present the accuracy of the global networking (direct-physical and indirect-functional interactions) (Szklarczyk et al., 2019, 2023). Zenda et al. (2018), performed PPI network analysis of DAPs (Differentially Abundant Proteins) from drought stressed two varieties of maize seedling leaves (YE8112 and MO17). The several groups of interacting proteins obtained linked to amino acid metabolism, energy metabolism (NADPH production), electron transport and stress signalling, and maintaining redox homeostasis with confidence score higher than 0.5. The network analysis of DEGs (Differentially expressed genes)of Hordeum vulgare presented the networking interconnection related to metabolic processes, cellular process, localization and biological regulation(Zinati et al., 2021). Similarly, protein–protein interaction (PPI)analysis of DEGs of several plant species leaves yielded some key upregulatory and downregulatory proteins in response to drought, cold and heat stress (Balti et al., 2021)

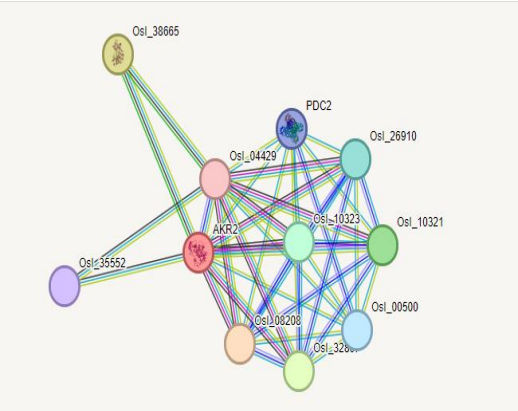

Figure 5: Protein Network and Enrichment Analysis of AKR4C14obtained from String.

4. Functional Analysis of AKR4C14

The functional analysis performed by Interproscan results provide the information related to several important motifs present in AKR4C14 protein. The largest region belongs to NADP-dependent oxidoreductase domain and AKR family (Figure 7).The functional analysis or gene annotation of AKR4C14 (6KBL)protein sequence is summarized in the Table 3.

This analysis reveals the key protein involved in essential biological process such as aroma

production, stress resistant quality that our characteristics of Thai Jasmine Rice. The similar results were obtained for Solanum lycopersicum which displayed atleast one AKR domain in all 28 members of AKR genes(Guan et al., 2023).Shalev et al. (2018)identified terpene synthases of Thujaplicata covering 2976 identified domains across 2904 InterPro protein

.

Figure 6: Protein Predicted Functional Partners Analysis Using String

Table 3: Gene annotation of AKR4C14 protein sequence

Category	Description	Residue Range (Amino Acids)
Representative Domains	AKR4C1-15 domain	6-292
Family	Aldo-keto reductase (AKR)	2-309
Domain	NADP-dependent oxidoreductase	20-285
Homologous Superfamily	NADP-dependent oxidoreductase domain	4-314
Conserve Site	Aldo/keto reductase	42-59, 143-160, 254-269
Others	NADPH-dependentaldo-keto reductase,chloroplastic	5-314
Active Sites (Catalytic Tetrad)	Serine, Histidine, Aspartate Family	Some wider range residues starting at: 22 (G), 114 (H), 159 (N), 206 (Y), 254 (L), and 265 (Q)

families, the largest of which being the family of cytochrome P450s with 565 sequences. The 50 probesets of myo-inositol O-methyltransferase (MtIMT) identified by Interpro scan to obtain proteins with NAD(P)(H)-binding domains that could be potential OEP (D-ononitol epimerization pathway) enzymes in Medicago truncatula(Pupel et al., 2019).

5. Functional Motif or Domain Analysis

The ScanProsite analysis of AKR4C14protein sequence interpreted two notable hits, indicating that our protein sequence matched two distinct motifs from the PROSITE database. The outcome is summarized in Table 4.

These protein signature sequence analysis, which may provide insights into the protein's biological roles and potential interactions. Both the hits obtained confidence level zero which represents reliable cut-off and high confidence level,reporting significant matches of the sequence (Figure 8).

Figure 7: Identifications of AKR4C14Protein Domain Families Functional Site.

Table 4: Description of notable hits.

Chain	Motif	Start	Stop	Position of first amino acid in structure === Match sequence	Colour
A	PS00063-ALDOKETO_REDUCTASE_3	254	269	1 === LPKSVSEERIKQNLDV	Red
A	PS00798-ALDOKETO_REDUCTASE_1	42	59	1 === GYRHIDCASAYNNEKEVG	Mustard

Protein function conserved motifs or cofactor binding sites are found using the ScanProsite (Robinson et al., 2021). The Meyerozyma caribbica sequence oxidoreductase activity was found to have substantial biological sites by Arumugam et al. (2020). The query coverage was 100%, and the e-value of 9e-86 corresponded to the XR of Candida tenuis xylose reductase, which has Protein Data Bank (PDB) id 1JEZ. Similarly, this tool's study of the 1,3,6,8 tetra hydroxynaphthalene reductase protein sequence FOXG_04696 (T4HNR like belongs to Fusarium oxysporum, vascular wilt of tomato) indicated that the sequence has signature sequences that belong to the SDR (short-chain dehydrogenases/reductases) family. This enzyme (SDRs) has a role in redox sensing and catalyses the

metabolism of organic macromolecules, including cofactors, carbohydrates, lipids, amino acids, steroids, and aromatic chemicals (Aamir et al.,2018)

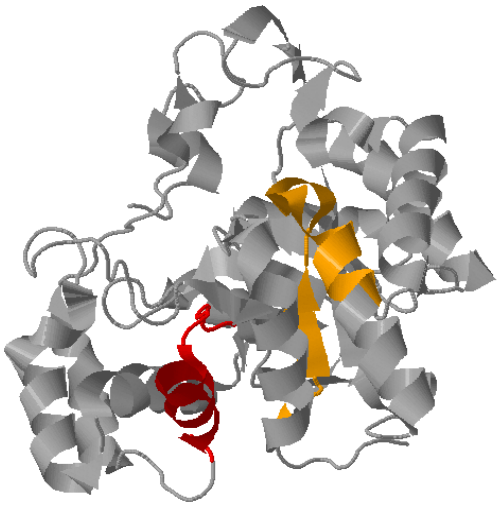

Figure 8: Notable two hits (coloured region) by ScanProsite Analysis of AKR4C14

6. NGS with Biopython Analysis

The biopython results presented the amino acid residue positions of 6KBL protein in the form of scatter plot graph (Figure 8). The code also extracts the alignment of sequences from a file and calculates pairwise sequence similarities using the BLOSUM62 substitution matrix (Figure 9).The GRAVY of the sequence 6KBLwas calculated as -0.227, which depicts the hydrophilic nature of protein. The coding also made possible to calculate each amino acid residue quantitatively and qualitatively (Figure 10).

```
plt.figure(figsize=(10, 6))
        plt.scatter(residue_positions, chain_ids, alpha=0.6)
        plt.xlabel('Residue Position')
        plt.ylabel('Chain ID')
        plt.title('Residue Positions in PDB Structure 6kbl')
        plt.grid(True)
        plt.show()
```

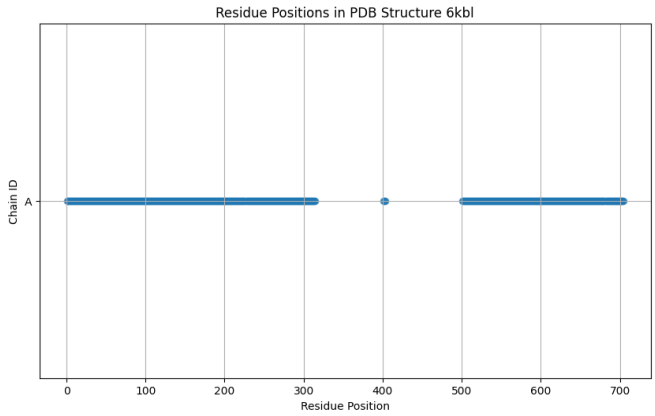

Figure 8: Excessive residue position in PDB structure using Biopython

```
print("Similarity Matrix:")

print(similarity_matrix)
```

```
Similarity Matrix:
[[5.02484472 4.36335404 4.46583851]
 [4.36335404 4.9378882 4.79192547]
 [4.46583851 4.79192547 5.28571429]]
```

Figure 9: Similarity Scores

```
Amino Acid Composition for first chain:
A: 0.09%
C: 0.01%
D: 0.05%
E: 0.06%
F: 0.03%
G: 0.07%
H: 0.03%
I: 0.06%
K: 0.08%
L: 0.11%
M: 0.02%
N: 0.04%
P: 0.06%
Q: 0.04%
R: 0.03%
S: 0.08%
T: 0.03%
V: 0.07%
W: 0.03%
Y: 0.03%
```

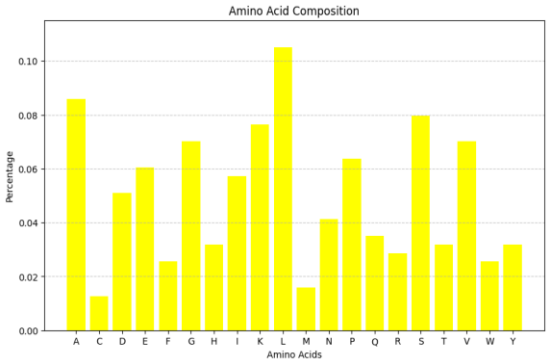

Figure 10: Performing an amino acid composition analysis and visualizing the result using a graph in biopython

The comprehensive proteomics analysis combined with NGS and Biopython tools provided valuable insights into the functional roles of AKR4C14. The enzyme's involvement in stress response and its regulatory mechanisms underscore its importance in rice physiology. Future studies could focus on functional validation through gene knockout or overexpression experiments to further elucidate AKR4C14's roles in rice.Additionally, a hydrophilic amino acid score of -0.227 suggests functional implications. Overall, this study integrates next-generation sequencing (NGS) and Biopython for a detailed proteomics analysis of the AKR4C14 enzyme, emphasizing its potential role in stress tolerance and setting the stage for further research on metabolic regulation in rice

Conclusion

This analysis identified specific proteins enriched in Jasmine Rice, highlighting their potential importance for breeding. Future experimental studies could validate these findings, with InterPro scans aiding rice genetics and biotechnology applications. Insights from multiple sequence alignment (MSA) will enhance understanding of molecular evolution, while Biopython's bio.PDB module facilitates structural analysis,genetic analysis, molecular docking analysis, protein structure analysis.Jasmine rice analysis using Next generation sequencing (NGS) data and Biopython to analyze jasmine rice,the conclusion and potential outcomes for genomic diversity and improvement,identifying important traits related to yield disease resistance and aroma.In functional proteomics,Evolutionary studies and rice varieties comparisions.Overall, the integration of NGS and Biopython for Jasmine rice research holds tremendous potential for future development.

References

1. Aamir, M., Singh, V. K., Dubey, M. K., Meena, M., Kashyap, S. P., Katari, S. K., Upadhyay, R. S., Umamaheswari, A., & Singh, S. (2018). In silico Prediction, Characterization, Molecular Docking, and Dynamic Studies on Fungal SDRs as Novel Targets for Searching Potential Fungicides Against Fusarium Wilt in Tomato. Frontiers in Pharmacology, 9. https://doi.org/10.3389/fphar.2018.01038
2. Arumugam, N., Boobalan, T., Saravanan, S., Jothi Basu, M., Arun, A., Suganya Devi, T., & Kavitha, T. (2020). In silico and in vitro comparison of nicotinamide adenine dinucleotide phosphate dependent xylose reductase rossmaan fold in Debaryomycetaceae yeast family. Biocatalysis and Agricultural Biotechnology, 24, 101508. https://doi.org/10.1016/j.bcab.2020.101508
3. Ashraf, M. F., Hou, D., Hussain, Q., Imran, M., Pei, J., Ali, M., Shehzad, A., Anwar, M., Noman, A., Waseem, M., & Lin, X. (2022). Entailing the Next-Generation Sequencing and Metabolome for Sustainable Agriculture by Improving Plant Tolerance. International Journal of Molecular Sciences, 23(2), 651. https://doi.org/10.3390/ijms23020651
4. Balti, I., Benny, J., Perrone, A., Caruso, T., Abdallah, D., Salhi-Hannachi, A., & Martinelli, F. (2021). Identification of conserved genes linked to responses to abiotic stresses in leaves among different plant species. Functional Plant Biology, 48(1), 54. https://doi.org/10.1071/FP20028
5. Fang, J., Zhu, X., Wang, C., &Shangguan, L. (2016). Applications of DNA Technologies in Agriculture. Current Genomics, 17(4), 379–386. https://doi.org/10.2174/1389202917666160331203224
6. Guan, X., Yu, L., & Wang, A. (2023). Genome-Wide Identification and Characterization of Aldo-Keto Reductase (AKR) Gene Family in Response to Abiotic Stresses in Solanum lycopersicum. International Journal of

Molecular Sciences, 24(2), 1272. https://doi.org/10.3390/ijms24021272
7. Pupel, P., Szablińska-Piernik, J., &Lahuta, L. B. (2019). Two-step <scp>d</scp> -ononitol epimerization pathway in Medicago truncatula. The Plant Journal, 100(2), 237–250. https://doi.org/10.1111/tpj.14439
8. Ramanathan, V., Rahman, H., Subramanian, S., Nallathambi, J., Kaliyaperumal, A., Manickam, S., Ranganathan, C., &Muthurajan, R. (2018). OsARD4 encoding an acireductone dioxygenase improves root architecture in rice by promoting development of secondary roots. Scientific Reports, 8(1), 15713. https://doi.org/10.1038/s41598-018-34053-y
9. Robinson, S. L., Piel, J., &Sunagawa, S. (2021). A roadmap for metagenomic enzyme discovery. Natural Product Reports, 38(11), 1994–2023. https://doi.org/10.1039/D1NP00006C
10. Sachan, S., Kumari, U., Naveen Kumar, M., &Dokka, M. K. (2024). Drug Discovery and Homology Modeling Study of AKR4C14, an Aldo-Keto Reductase from Thai Jasmine Rice. https://doi.org/10.48047/AFJBS.6.14.2024.8971-8984
11. Samal, K. C., Sahoo, J. P., Behera, L., & Dash, T. (2021). Understanding the BLAST (Basic Local Alignment Search Tool) Program and a Step-by-step Guide for its use in Life Science Research. Bhartiya Krishi AnusandhanPatrika, Of. https://doi.org/10.18805/BKAP283
12. Sameer, H., Victor, G., Katalin, S., & Henrik, A. (2021). Elucidation of ligand binding and dimerization of <scp>NADPH</scp> :protochlorophyllide (Pchlide) oxidoreductase from pea (<scp>Pisum sativum</scp> L.) by structural analysis and simulations. Proteins: Structure, Function, and Bioinformatics, 89(10), 1300–1314. https://doi.org/10.1002/prot.26151
13. Shalev, T. J., Yuen, M. M. S., Gesell, A., Yuen, A., Russell, J. H., &Bohlmann, J. (2018). An annotated transcriptome of highly inbred Thujaplicata (Cupressaceae) and its utility for gene discovery of terpenoid biosynthesis and conifer defense. Tree Genetics & Genomes, 14(3), 35. https://doi.org/10.1007/s11295-018-1248-y
14. Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. Protein Science, 27(1), 135–145. https://doi.org/10.1002/pro.3290
15. Songsiriritthigul, C., Narawongsanont, R., Tantitadapitak, C., Guan, H.-H., & Chen, C.-J. (2020). Structure–function study of AKR4C14, an aldo-keto reductase from Thai jasmine rice (Oryza sativa L. ssp. indica cv. KDML105). Acta Crystallographica Section D Structural Biology, 76(5), 472–483. https://doi.org/10.1107/S2059798320004313
16. Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., &Mering, C. von. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Research, 47(D1), D607–D613. https://doi.org/10.1093/nar/gky1131
17. Szklarczyk, D., Kirsch, R., Koutrouli, M., Nastou, K., Mehryary, F., Hachilif, R., Gable, A. L., Fang, T., Doncheva, N. T., Pyysalo, S., Bork, P., Jensen, L. J., & von Mering, C. (2023). The STRING database in 2023: protein–protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Research, 51(D1), D638–D646. https://doi.org/10.1093/nar/gkac1000
18. Zinati, Z., Sazegari, S., Tahmasebi, A., &Delavari, A. (2021). A comprehensive meta-analysis to identify transcriptional signatures of abiotic stress responses in barley (Hordeum vulgare). Cereal Research Communications, 49(3), 385–391. https://doi.org/10.1007/s42976-020-00107-z