doi 10.18699/bgrs2024-2.2-16

Democratized Bioinformatics: Enabling accessibility through an integrated platform for drug discovery and multi-omics analysis

Mohammad T.[1, 2]*, Khan S.[2], Hassan M.I.[2]*

[1] *Department of Biophysics, All India Institute of Medical Sciences (AIIMS), New Delhi 110029, India*

[2] *Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi 110025, India*

** tajm@aiims.edu; mihassan@jmi.ac.in*

Keywords: Computational biology; Bioinformatics platform; Drug discovery; Multi-omics data analytics; Democratization of bioinformatics

Motivation and Aim: The rapidly evolving landscape of modern drug discovery and therapeutic management relies heavily on bioinformatics and computational biology tools. However, these resources often pose accessibility and affordability challenges, particularly for researchers with limited resources or expertise. To address these issues, there is a need for an integrated bioinformatics platform that combines state-of-the-art algorithms and AI models into a user-friendly interface, specifically tailored for modern drug discovery pipelines and multi-omics data analytics. This platform aims to democratize bioinformatics techniques, empower researchers across diverse scientific domains, and revolutionize the utilization of bioinformatics and computational biology.

Methods and Algorithms: The development of this platform involved a comprehensive methodology that includes requirement analysis, module development, data collection and preprocessing, module training and evaluation, platform development, testing and optimization, deployment and accessibility, validation and user feedback, dissemination and impact evaluation, and long-term maintenance and sustainability. Specialized modules focusing on various tasks such as virtual screening, molecular simulations, QSAR modeling, pharmacophore mapping, and NGS analytics have been developed and integrated into the platform using Python and R programming languages. Various AI models were trained and evaluated using diverse datasets to ensure accuracy and reliability.

Results: Several specialized programs have already been developed, deployed, and published [1, 2]. These programs perform tasks such as single-click high-throughput virtual screening and sequence and structure-based protein analyses. Currently, multiple programs/bioinformatics tools and algorithms are being developed and integrated into an integrated platform. The anticipated outcomes include the development of an integrated bioinformatics platform, specialized modules, benchmarked algorithms, an open-access platform, and case studies showcasing the platform's practical applications in drug discovery and multi-omics analyses. These outcomes aim to enhance research capability, democratize bioinformatics, contribute to national missions in biotechnology and healthcare, accelerate drug discovery, enable data-driven healthcare, foster scientific collaboration, have an educational impact, and increase publication and visibility.

Conclusion: This project addresses critical challenges in modern drug discovery and computational biology by developing an integrated bioinformatics platform that is accessible, user-friendly, and tailored for diverse research needs. By democratizing bioinformatics techniques and empowering researchers, this platform has the potential to accelerate scientific discoveries and foster collaborative research endeavors worldwide.

Acknowledgment: Taj Mohammad acknowledges SERB National Post Doctoral Fellowship (NPDF) with reference number: PDF/2022/001793.

References

1. Mohammad T., Mathur Y., Hassan M.I. InstaDock: A single-click graphical user interface for molecular docking-based virtual high-throughput screening. *Brief Bioinform*. 2021;22(4):bbaa279
2. Mathur Y., Mohammad T., Anjum F. et al. PyPAn: An automated graphical user interface for protein sequence and structure analyses. *Prot Pept Lett*. 2022;29(4):306-12