Bioinformatics in Metabolic Engineering

Nida Tabassum Khan[*]**, Namra Jameel, Muhammad Jibran Khan**

Department of Biotechnology, Faculty of Life Sciences and Informatics,

Balochistan University of Information Technology Engineering and Management Sciences, (BUITEMS),

Quetta, Pakistan

*****Corresponding Author:** *Nida Tabassum Khan, Department of Biotechnology, Faculty of Life Sciences and Informatics, Balochistan University of Information Technology Engineering and Management Sciences, (BUITEMS), Quetta, Pakistan.*

Abstract: *The emergence of metabolic engineering is strengthened by the tools of bioinformatics that not only supports experimental techniques employed for the manipulation of biological pathways but also provides an insilco platform for their study. There are numerous bioinformatics tools and software such as KEGG, MetaCyc,PH USER, Cytoscape etc that could be utilized by researchers in effectively transforming cellular pathways.*

Keywords: *Systems biology ontology; Metabolic pathways; Gene network; Network visualization*

1. INTRODUCTION

Metabolic engineering integrates efficient exploration of metabolic and cellular pathways with molecular systems to enhance biological properties by formulating and applying genetic alterations [1].Decades ago, computational tools were used in metabolic engineering to design single enzymes by using analog computers, and digital computers were used to design the dynamics of metabolic pathways [2]. In recent years, bioinformatics tools are used to search huge metabolic databases leading to better understanding, managing, analyzing and visualizing of metabolic pathways [3, 4].

2. APPLICATIONS OF BIOINFORMATICS TOOLS IN METABOLIC ENGINEERING

There are several applications of bioinformatics tools in metabolic engineering which are as follows:

2.1. Metabolic Networks Reconstructing

Reconstructions of metabolic networks identifies the crucial components that are of prime functional importance in a biological pathway of a particular organism by linking data obtained from metabolic and genomic studies [5]. The measure of metabolic reconstructions may lie from studying distinct pathways to investigating whole genomes [6]. Metabolic network size and biological databases are potent factors that influences the effectiveness of network reconstruction [7].Analyzing and designing sequences, sorting and retrieving metabolic data, and representing network information are the significant errands related with the reconstruction of these biological networks[8].Whether metabolic network reconstruction is achieved manually or by employing bioinformatics tools, it is the critical prime phase in any metabolic engineering application [9].

Examples of some soft wares used for metabolic network construction are given below:

- KAAS or PathoLogic: These two softwares are commonly used to program a reconstruction [10].

- KEGG and MetaCyc: Kyoto Encyclopedia of Genes and Genomes (KEGG) and METACYC both exhibits metabolic networks as a map that depicts series of reactions that expresses interactions between biological catalysts and their associated substrates [11].

- SBML, SBO and BioPAX: Systems Biology Markup Language (SBML) is used to represent metabolic pathways models [12] and Systems Biology Ontology (SBO) is used to stipulate the sense of parameters present in the kinetic models of specific pathway [13].While Biological

Pathway Exchange Language (Bio PAX) permits biological databases and tools to exchange data concerning the qualitative functional associations of networks [14].

2.2. Metabolic Networks visualizing

Metabolic networks are basically pathways depicting biological catalysts assisted reactions that involve the conversion of one metabolite in to another [15].Visualizing such reactions aids in better apprehension and identification of parameters that influence pathway effectiveness [16].

Examples of some soft wares used for metabolic network visualizing are given below:

- Cytoscape, GraphViz and Systrip: These are network visualization tools that functions on algorithms in order to interpret large interaction networks [17, 18].

- Systems Biology Graphical Notation (SBGN): SBGN is utilized for representing biochemical network illustrations [19].

- Synthetic Biology Open Language (SBOL): SBOL is a network visualization tool that precisely identifies the major constituents in a gene network diagram, and can be used to recreate a gene network using information generated in the knowledgebase [20].

2.3. Manipulating Biological Molecules

Synthesizing DNAs, RNAs, proteins or enzymes, constructing translational cassettes, and mutating genomic sequences requires physical alterations of biological entities [21]. Bioinformatic tools relates sequence data with experimentally determined morphological or functional states of a biological molecule for designing molecules with enhanced functionality[22].

Examples of some soft ware's used for manipulating biological molecules

- Tinker Cell and Geno CAD: These two in silco tools streamlines the metabolic process by integrating enzymes into functional circuits to produce a diagram that identifies gene function by using a set of symbols that represents components such as promoters, coding sequences or terminators [23,24].

- Geneious: It is a sequence editing tool with significant visualization, organization, assemblage, and alignment features to identify features such as allosteric sites, promoter, binding sites, origin of replication etc[25, 26].

- Primer help for user(PHUSER): It is user friendly software that aids user in designing primers for the molecular sequences [27].

- Gene Designer: This computational tool can suggest a DNA sequence once users have provided either a DNA or protein sequence as a query and a cod on usage table for a certain organism [28]

2.4. Metabolic Flux Analysis

Metabolic flux analysis (MFA) determines the metabolic fluxes in vivo with the aim of accelerating product synthesis [29].

Examples of some soft wares used for metabolic flux analysis are as follows:

- FBA, MOMA and ROOM: These tools assist in targeted manipulations that lead to a desired phenotype [30].

- COBRA: Constraint based reconstruction and analysis (COBRA) methods have been employed in transcriptional modeling and signaling networks [31]. These methods have been used to aid metabolic pathway engineering, to model antigens and host-pathogen interactions and to measure the impact of disease states on human metabolism [32].

2.5. Pathway Prospecting\Searching

Following are some bioinformatics tools used for pathway prospecting in metabolic engineering

- Opt Strain: It is an algorithm that quests numerous databases to explore, evaluate, and select pathways that lead to the fabrication of a target molecule [33].

- Opt Knock: To further enhance target production through targeted genomic alteration Opt Knock is employed [34].

3. CONCLUSION

Collectively, these computational bioinformatics tools provide an influential background to enhance design and analysis aspects of metabolic engineering. Thus plays a significant role in metabolic engineering that can be employed in diverse applications such as in the production of chemicals, fuels etc., primer designing, drug designing etc.

REFERENCES

[1] Bailey, J. E. (1991). Toward a science of metabolic engineering. Science, 252(5013), 1668-1675.

[2] Nielsen, J. (2001). Metabolic engineering. Applied microbiology and biotechnology, 55(3), 263-283.

[3] Stephanopoulos, G., Aristidou, A. A., & Nielsen, J. (1998). Metabolic engineering: principles and methodologies. Elsevier.

[4] Keasling, J. D. (2010). Manufacturing molecules through metabolic engineering. Science, 330(6009), 1355-1358.

[5] Blazier, A. S., & Papin, J. A. (2012). Integration of expression data in genome-scale metabolic network reconstructions. Frontiers in physiology, 3, 299.

[6] Henry, C. S., Bernstein, H. C., Weisenhorn, P., Taylor, R. C., Lee, J. Y., Zucker, J., & Song, H. S. (2016). Microbial community metabolic modeling: a community data-driven network reconstruction. Journal of cellular physiology, 231(11), 2339-2345.

[7] Chen, N., del Val, I. J., Kyriakopoulos, S., Polizzi, K. M., & Kontoravdi, C. (2012). Metabolic network reconstruction: advances in in silico interpretation of analytical information. Current opinion in biotechnology, 23(1), 77-82.

[8] Manichaikul, A., Ghamsari, L., Hom, E. F., Lin, C., Murray, R. R., Chang, R. L., ... & Thiele, I. (2009). Metabolic network analysis integrated with transcript verification for sequenced genomes. Nature methods, 6(8), 589.

[9] Heavner, B. D., & Price, N. D. (2015). Transparency in metabolic network reconstruction enables scalable biological discovery. Current opinion in biotechnology, 34, 105-109.

[10] Pitkänen, E., Rousu, J., & Ukkonen, E. (2010). Computational methods for metabolic reconstruction. Current opinion in biotechnology, 21(1), 70-77.

[11] Hamilton, J. J., & Reed, J. L. (2014). Software platforms to facilitate reconstructing genome-scale metabolic networks. Environmental microbiology, 16(1), 49-59.

[12] Ghosh, S., Matsuoka, Y., Asai, Y., Hsin, K. Y., & Kitano, H. (2011). Software for systems biology: from tools to integrated platforms. Nature Reviews Genetics, 12(12), 821.

[13] Hoehndorf, R., Dumontier, M., Gennari, J. H., Wimalaratne, S., de Bono, B., Cook, D. L., & Gkoutos, G. V. (2011). Integrating systems biology models and biomedical ontologies. BMC systems biology, 5(1), 124.

[14] Strömbäck, L., & Lambrix, P. (2005). Representations of molecular pathways: an evaluation of SBML, PSI MI and BioPAX. Bioinformatics, 21(24), 4401-4407.

[15] Droste, P., Miebach, S., Niedenführ, S., Wiechert, W., & Nöh, K. (2011). Visualizing multi-omics data in metabolic networks with the software Omix—a case study. Biosystems, 105(2), 154-161.

[16] Copeland, W. B., Bartley, B. A., Chandran, D., Galdzicki, M., Kim, K. H., Sleight, S. C., ... & Sauro, H. M. (2012). Computational tools for metabolic engineering. Metabolic engineering, 14(3), 270-280.

[17] Jourdan, F. (2013). Qualitative Modelling of Metabolic Networks. In Advances in Botanical Research (Vol. 67, pp. 557-591). Academic Press.

[18] Auber, D., Archambault, D., Bourqui, R., Lambert, A., Mathiaut, M., Mary, P., ... & Mélançon, G. (2012). The tulip 3 framework: A scalable software library for information visualization applications based on relational data.

[19] Le Novere, N., Hucka, M., Mi, H., Moodie, S., Schreiber, F., Sorokin, A., ... & Bergman, F. T. (2009). The systems biology graphical notation. Nature biotechnology, 27(8), 735.

[20] Bartley, B., Beal, J., Clancy, K., Misirli, G., Roehner, N., Oberortner, E., ... & Zhang, Z. (2015). Synthetic biology open language (SBOL) version 2.0. 0. Journal of integrative bioinformatics, 12(2), 902-991.

[21] Yadav, V. G., De Mey, M., Lim, C. G., Ajikumar, P. K., & Stephanopoulos, G. (2012). The future of metabolic engineering and synthetic biology: towards a systematic practice. Metabolic engineering, 14(3), 233-241.

[22] Keasling, J. D., & Chou, H. (2008). Metabolic engineering delivers next-generation biofuels. Nature biotechnology, 26(3), 298.

[23] Chandran, D., Bergmann, F. T., & Sauro, H. M. (2010). Computer-aided design of biological circuits using TinkerCell. Bioengineered bugs, 1(4), 276-283.

[24] Wilson, M. L., Hertzberg, R., Adam, L., & Peccoud, J. (2011). A step-by-step introduction to rule-based design of synthetic genetic constructs using GenoCAD. In Methods in enzymology (Vol. 498, pp. 173-188). Academic Press.

[25] Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., ... & Thierer, T. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics, 28(12), 1647-1649.

[26] Olsen, C., Qaadri, K., Moir, R., Kearse, M., Buxton, S., & Cheung, M. (2014). Geneious R7: a bioinformatics platform for biologists. In International Plant and Animal Genome Conference Xxii.

[27] Olsen, L. R., Hansen, N. B., Bonde, M. T., Genee, H. J., Holm, D. K., Carlsen, S., ... & Wernersson, R. (2011). PHUSER (Primer Help for USER): a novel tool for USER fusion primer design. Nucleic acids research, 39(suppl_2), W61-W67.

[28] Villalobos, A., Ness, J. E., Gustafsson, C., Minshull, J., & Govindarajan, S. (2006). Gene Designer: a synthetic biology tool for constructing artificial DNA segments. BMC bioinformatics, 7(1), 285.

[29] Bonarius, H. P., Schmid, G., & Tramper, J. (1997). Flux analysis of underdetermined metabolic networks: the quest for the missing constraints. Trends in Biotechnology, 15(8), 308-314.

[30] Zhao, Q., & Kurata, H. Comparison of the Prediction Abilities of FBA, MOMA and ROOM for a pykF Mutant of E. coli. In The 17th International Conference on Genome Informatics (pp. 107-1).

[31] Schellenberger, J., Que, R., Fleming, R. M., Thiele, I., Orth, J. D., Feist, A. M., ... & Kang, J. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2. 0. Nature protocols, 6(9), 1290.

[32] Hyduke, D., Schellenberger, J., Que, R., Fleming, R., Thiele, I., Orth, J., ... & Rahmanian, S. (2011). COBRA Toolbox 2.0. Protocol exchange, 22.

[33] Pharkya, P., Burgard, A. P., & Maranas, C. D. (2004). OptStrain: a computational framework for redesign of microbial production systems. Genome research, 14(11), 2367-2376.

[34] Burgard, A. P., Pharkya, P., & Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnology and bioengineering, 84(6), 647-657.