Int.J.Curr.Microbiol.App.Sci (2024) *13*(01): 94-103

International Journal of Current Microbiology and Applied Sciences
ISSN: 2319-7706 **Volume 13 Number 1 (2024)**
Journal homepage: http://www.ijcmas.com

Review Article

https://doi.org/10.20546/ijcmas.2024.1301.012

An OMICS-Based Approach Studies Natural Products

P. A. Nimitha*

St.Raphaels CGHSS Ollur, Thrissur, Kerala, India

**Corresponding author*

Keywords

Bioinformatics, Omics, Plant studies, Proteomics, Transcriptomics

Article Info

Received:
25 November 2023
Accepted:
26 December 2023
Available Online:
10 January 2024

A B S T R A C T

In conjunction with bioinformatics and comparable developments in tools, software, and visualisation modelling, current developments in plant sciences have propelled the scientific community into an active dispute over information. Despite the advent of Omics and numerous other remarkable bioinformatics tools, a considerable proportion of researchers still require further familiarisation with these instruments. The present evaluation centres on the potential implementations of diverse in silico tools and technologies in the analysis of plant sciences. Gaining knowledge of these many technologies will contribute to an enhanced comprehension of plant characteristics, including resistance to pathogens, tolerance to stress, and nutritional enhancement. Furthermore, we are collaborating on many challenges and limitations in the field of plant sciences that are associated with the bioinformatics methodology.

Introduction

Natural products have perpetually occupied a pivotal role in human history and advancement by functioning as the primary source of medicine and food (Vadakkan *et al.,* 2018; Kayeen Vadakkan *et al.,* 2018c; Vadakkan *et al.,* 2019a). It occasionally provided protection, nutrition, and therapeutic resources for the human population. To date, numerous actions, encompassing both medicinal and non-medical aspects, have been investigated. These actions include larvicidal, anti-cancer, anti-helminthic, and antibacterial properties (Kayeen Vadakkan *et al.,* 2018a; Vadakkan, 2019; Vadakkan *et al.,* 2019c, 2019b). In tandem with the heightened importance placed on plants, the field of plant research has produced a tremendous amount of data. The rapid expansion of digital data is a consequence of recent technological and hardware developments that enable the analysis of biological samples on a nanoscale and macroscopic levels. Due to the inability of the human brain to handle this vast amount of information, there is a rapidly growing need for computer methods to analyse and contextualise it. Bioinformatics is the application of principles and methodologies derived from the fields of engineering, statistics, and computer science to the study of biological data (Kayeen Vadakkan *et al.,* 2018b, 2018d; Vadakkan, 2020; Vadakkan *et al.,* 2021). It finds its greatest applicability in the domains of computational biology and biological information management.

Bioinformatics is currently being utilised in the field of computational biology for a multitude of significant objectives that extend beyond the mere analysis of genomic sequence data. The aforementioned tasks encompass the examination of gene expression and variation, the configuration of simulation settings, the

identification and prediction of gene regulatory networks, and the analysis and modelling of the structure and function of proteins and genes (Suriyakala *et al.,* 2021; Vadakkan *et al.,* 2020).

Transcriptomics-Understanding the expressed ones

In recent years, transcriptomics has emerged as one of the most actively researched topics in the field of biology. RNA-Seq is a method that may be used to research model organisms as well as non-model species because of its high throughput, high sensitivity, and high resolution. The sequencing of transcriptomes is another method that is crucial for researching the genomes of plants, a topic for which there is a paucity of information. (Giacomello, 2021; Guo *et al.,* 2021). Because of advancements in technology, transcriptomics has been an increasingly popular area of research since the late 1990s. In the early 1990s, researchers made their initial attempts to investigate the whole transcriptome. The field of transcriptomics has been moulded in many ways by the many advances that have been made in technology. RNA sequencing, which makes use of high-throughput sequencing to capture all sequences, and microarrays, which measure a collection of pre-set sequences, are two key modern methodologies in the field of study. (Chen *et al.,* 2021; Chu *et al.,* 2017; Libault *et al.,* 2017). Transcriptomics and genomes are distinguished from one another by a number of factors. First, transcriptomic analysis may be used to track an organism's overall transcriptional activity without a reference genome, whereas genome assembly is more difficult and expensive in research. Second, the transcriptome changes based on the time and location of observation because it contains information on additional metabolic pathways in addition to reflecting variations in gene expression at various temporal and spatial sites (Barrera-Redondo *et al.,* 2020; Lowe *et al.,* 2017; Pierlé *et al.,* 2012)

The fundamental objective of doing a transcriptome study is to get an understanding of how variations in the quantity of transcripts influence the growth and development of an organism as well as its reaction to its surrounding environment. DNA microarrays have shown to be an effective technique for analysing the transcriptional profile of genes located all throughout the genome. (Malone and Oliver, 2011; Rao *et al.,* 2019). Microarray analysis provides the simultaneous evaluation of transcript abundance for thousands of genes. There are two distinct types of microarrays: arrays with a large number of tiny probes produced directly and arrays amplified by polymerase chain reaction (Bumgarner, 2013). Large microarray data sets can be analysed in a number of ways using a range of technologies. Examples include Gene Traffic, Gene Spring, and the Gene Chip Operating Software from Affymetrix (GCOS) (Kumari *et al*., 2007; Ragoussis and Elvidge, 2006). Various tools involved in transcriptomic analysis are summarized in figure 2.

Transcriptomic sequence assembly

The presence of undesired rRNA reads is a common issue in RNA-Seq experiments; consequently, quality control is the first and essential step (DeLuca *et al.,* 2012; Kumar *et al.,* 2020; Zhou *et al.,* 2018). RNA-Seq data can be subjected to a preliminary quality check using the FastQC software. RNA-Seq data reports from FastQC should be read considerably differently from genomic data reports. RNA-Seq data are distinguished by their unequal coverage depth and the presence of ribosomal RNA sequences. rRNA readings may add new peaks into the GC content plot due to their high abundance (de Sena Brandine and Smith, 2019; Leggett *et al.,* 2013). After verifying that the quality is acceptable, we can move on to the sequence alignment step.

While the algorithms and techniques used for DNA sequence alignment may also be used to align bacterial RNA sequences to the reference genome. Due to the presence of splicing events, the mapping of eukaryotic RNA is drastically different. Algorithms for splicing alignment are needed in order to support lengthy insertions and avoid intronic sequences. The most widely used software programmes for mapping RNA to DNA are BLAT, GMAP, and Splign (Bhagwat *et al.,* 2012; Jammali *et al.,* 2019; Kapustin *et al.,* 2008; Wu and Watanabe, 2005)

Due to the lack of lengthy complicated repeats and the shorter total length of the transcriptome, the de novo transcriptome assembly may appear to be an easier task to solve than genome assembly (Hina *et al.,* 2020; Mahmood *et al.,* 2020; Raghavan *et al.,* 2022). The process of assembling the transcriptome is difficult because of the vast amount of data that is produced as a result of the varying levels of expression, the inclusion of isoforms produced by the same gene, and the presence of paralogous genes. The majority of current de Bruijn graph genome assemblers constitute the foundation for modern de novo transcriptome assembly tools like Trans-

ABySS and SOAPdenovo-Trans (Sessegolo *et al.,* 2019; Simpson *et al.,* 2009; Yang *et al.,* 2019; Yang and Smith, 2013).

Bioinformatics tools for analysing transcriptomic data

The initial phase of a conventional transcriptomic investigation is data generation. Conventionally, mRNAs or whole RNAs are transformed into a cDNA library, which is then fragmented and sequenced to produce single-end or paired-end short reads (Conesa *et al.,* 2016; Wang *et al.,* 2019; Yang and Kim, 2015). Transcriptomic data analysis poses difficulties as the number of transcripts to rebuild is unknown, As RNA-seq can measure expression levels across 5 orders of magnitude, any gene can be studied by hundreds of thousands or a few reads. Even within gene or transcript constraints, library preparation and sequencing bias causes variable read coverage (Martin *et al.,* 2013; Wang *et al.,* 2009). Regarding transcript assembly, there are two different approaches: de novo, which assembles reads solely based on sequence overlap; and genome-based, which first aligns reads to a reference genome before assembling the alignments that overlap. In general, methods based on the genome are more accurate. In the absence of a genome sequence or in the event of a severely fragmented genome, de novo assembly, on the other hand, may be employed to create a representative collection of transcripts (Geniza and Jaiswal, 2017; Moreno-Santillán *et al.,* 2019; Ungaro *et al.,* 2017).

Assembling is more difficult and error-prone, yet it is done with the aid of some potent tools. Some very commonly used tools are discussed below. TopHat is an open-source program for aligning RNA-Seq reads to a reference genome beyond utilizing or depending on the known splice sites. TopHat aligns RNA-Seqreads taken in FASTA or FASTQ format using a reference genome (Kim *et al.,* 2013; Trapnell *et al.,* 2009). Initially unmapped reads (IUM reads) are the reads that have been placed aside because they do not map to the genome. Tophat examines the IUM reads for finding reads that pan junctions for each splice junction. For consensus-based building of mapped areas, the MAQ assembly module is utilised. TopHat is available for Mac OS X and Linux and is written in C++. It utilises the MAQ and Bowtie programmes and the SeqAn library (Spies and Ciaudo, 2015). Cufflinks Assembler is a different open-source C++ software that runs on Linux and Mac OS X. It could recognise whole novel transcripts and deterministically assign reads to isoforms. The Cuffdiff and Cuffcompare utilities are also included. Cuff comparison, coupled with transfrags (assembled transcript fragments), to annotated transcriptomes and detection of transfrags that are prevalent across several assemblies, are used to verify Cufflinks output (Babarinde *et al.,* 2019; Ghosh and Chan, 2016; Trapnell *et al.,* 2010). GFOLD, a fold change algorithm, creates biologically relevant rankings of differentially expressed genes from RNA-Seq data. According to the posterior distribution of logfold change, GFOLD provides reliable expression statistics when used to single-replica data sets (Feng *et al.,* 2012). The edgeR is a statistical method for profiling differential gene expression that is based on the negative binomial distribution. Although the edgeR was designed to work with replicates, it can also be applied to data sets without repetitions (Li *et al.,* 2022; Robinson *et al.,* 2010; Squair *et al.,* 2021).

Proteomic analysis-Third layer of Omics studies

Plants are continually exposed to a wide array of challenges and stresses in the natural environment in which they grow, which even pose a risk to the existence of spices. Plants have developed a variety of molecular programmes during the course of their evolution as a direct response to the changing environments that they have been subjected to. These programmes provide plants the ability to instantly detect and adjust to changes in their environment. Proteins are an essential component of plant response due to the fact that they play a direct part in the generation of novel plant phenotypes and are also accountable for the maintenance of cellular homeostasis (Hu *et al.,* 2015; Kosová *et al.,* 2018; Liu *et al.,* 2019). A protein's function is based not only on its molecular structure but also on its subcellular distribution and the modifications that occur after it is translated into a molecule. Protein function and subcellular distribution have a strong link due to the fact that different cell components provide different physiological and biochemical conditions. Alterations in the subcellular localization of proteins are a component of the vast majority of cellular biological processes and pathways. In light of this, the scientific world has been more interested in the study of proteomics. Recently (Cánovas *et al.,* 2004; Mergner and Kuster, 2022; Patole and Bindschedler, 2019; Winck *et al.,* 2021). Proteomics identifies and counts all protein types in a cell or tissue, investigates post-translational changes and protein interactions, and reveals protein molecules' structural and functional properties (Deswal *et al.,* 2013; Smythers and

Hicks, 2021). As a consequence of the fast development of sample pre-treatment and MS-based proteomic technology, qualitative proteome analysis is becoming more accurate, delivering expanded coverage and consistent quality. In the past, qualitative proteome analysis was limited to only finding proteins. The identification of proteins is no longer the primary emphasis; instead, an accurate and reliable quantitative analysis is being conducted. Since mass spectrometry cannot accomplish quantification on its own, other methods have been developed in order to get quantification through the use of mass spectrometry. These methods may be divided into two categories, namely labelling-based quantification and label-free quantification, depending on how they determine quantities (Angel *et al.,* 2012; Beck and Geiger, 2022; Macklin *et al.,* 2020; Rajczewski *et al.,* 2022).

Protein Sequencing and its analysis

The process of finding the sequence of amino acids, which may be written as either a single- or a three-letter code, is referred to as protein sequencing. The methods of protein sequencing may be broken down into one of two categories: methods that only produce the N-terminus sequence of a protein, and methods that sequence and identify the entire protein. (Alfaro *et al.,* 2021; Hunt *et al.,* 1986). Pehr Edman developed a label-cleavage method for protein sequencing in the 1950s. This method is based on a three-step reaction that labels and removes the N-terminal residue of a polypeptide. After this step, the polypeptide can be recognised as a phenylthiohydantoin (PTH – Edman reagent) derivative.

This technique is computer-aided and makes use of a protein sequencer in order to sequence peptides ranging in length from 5 to 50 amino acids. (Chen *et al.,* 2007; Miyashita *et al.,* 2001; Vecchi *et al.,* 2019; Walker, 1997). Fred Sanger developed a second method for conducting peptide end-group analysis. This method involves the utilisation of chemical derivatives to label the N-terminus of a protein with the yellow dye fluorodinitrobenzene, which is then followed by hydrolysis and the electrophoretic or chromatographic separation of the labelled N-terminal amino acid residue.

The whole amino acid sequence of a protein can be determined by doing many rounds of partial protein hydrolysis, fractionation, and determining the terminal amino acid (Callahan *et al.,* 2020; Rodriques *et al.,* 2019; Vitorino *et al.,* 2020).

Mass spectrometry (MS) methods are now the most common ones utilised for the sequencing and identification of proteins. The ratio of the masses of gas-phase ions is something that may be measured with an analytical technique called mass spectrometry. MS uses an electric current to disassemble peptides into their component amino acids, and then it gathers the freed amino acids in a mass spectrometer detector so that each can be identified according to the mass it possesses (Chen *et al.,* 2020; Dupree *et al.,* 2020; Han *et al.,* 2008; Standing, 2003). The invention of mass spectrometry technology made it feasible to sequence all of the proteins that are found in a live creature, which led to the birth of the field of study known as proteomics. The bottom-up strategy, which involves the study of peptide mixtures derived from digested proteins, and the top-down approach are the two primary types of mass spectrometry methods that are used for protein sequencing at the present time. The bottom-up method starts with proteolytic digestion of proteins to generate complex peptide samples, which are then analysed using high-throughput liquid chromatography and tandem mass spectrometry. On the other hand, the top-down method separates intact proteins from complex samples using liquid chromatography or 2-D gel electrophoresis. Both of these methods are used to study peptides. (de Graaf *et al.,* 2022; Hale, 2013; Pandeswari and Sabareesh, 2019; Singhal *et al.,* 2015; Tamara *et al.,* 2022).

Bioinformatics tools for Proteomic analysis

Because proteomic data consists of enormous quantity of data, it is important to use a broad variety of bioinformatic tools in order to analyse the data; some of the most common methods are described in the following paragraphs. PANTHER, which stands for Protein Analysis Through Evolutionary Relationships, is a classification system that can analyse sequencing, gene expression, and proteomics data. It combines ontology, gene function, pathways, and statistical methods (Karagiannis *et al.,* 2013; Mi *et al.,* 2013). A library and an index are both part of this exhaustive database of gene families, which was developed as a resource for the categorization of protein families and subfamilies. DAVID, which stands for Database for Annotation, Visualization, and Integrated Discovery, is yet another analytic tool used in this field. DAVID analyses enormous gene lists using the principle of gene enrichment to find genes that are functionally connected to a changed gene or protein (Dennis *et al.,* 2003a, 2003b; Hou *et al.,* 2022). The Kyoto Encyclopedia of

Genes and Genomes (KEGG) is a database resource that was developed for the purpose of analysing high-throughput data. It is separated into four categories, including information on systems, genomics, chemicals, and health. (Antonov *et al.,* 2008; Kanehisa and Goto, 2000; Xie *et al.,* 2017). Another software developed for analysis, understanding, integration and interpretation of biological data is Ingenuity Pathway Analysis (IPA) (Dong *et al.,* 2012; Yu *et al.,* 2016). Ingenuity another bioinformatic tool is required to analyse the data acquired from various platforms such as microarrays, proteomics, and metabolomics, among others. IPA makes use of the QIAGEN's Ingenuity Knowledge Base, which comprises materials that have been taken from articles, reviews, biological literature, and other sources and organised into Ontology terms. These materials may be found in the QIAGEN's Ingenuity Knowledge Base. In addition to this, a number of other tools are used, such as STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) and MINT, an open-source database of protein-protein interactions that is used to analyse protein networks. Both of these tools are utilised in the investigation of protein networks (Calderón-González *et al.,* 2016).

Since the field of plant sciences was first established, members of the scientific community have been striving to get a more in-depth understanding of the subject. At the beginning of the research process, it seemed like an impossible endeavour; nevertheless, thanks to advances in technology, we now have a better grasp of not just plants but all other living species as well. The invention of genomics, which made it possible to truly decode the gene pattern of an organism, was the initial step toward acquiring knowledge that led to the first enlightenment. Many people's ideas about how to comprehend an organism were fundamentally altered as a result of the development of transcriptomics and proteomics, which were supported by a wide variety of bioinformatics tools which was a significant step forward in the discipline.

Author Contribution

P. A. Nimitha: Investigation, formal analysis, writing—original draft.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Research Funding: Not applicable

Ethical Approval: Not applicable.

Consent to Participate: Not applicable.

Consent to Publish: Not applicable.

Conflict of Interest: The authors declare no competing interests.

References

Alfaro, J.A., Bohländer, P., Dai, M., Filius, M., Howard, C.J., van Kooten, X.F., Ohayon, S., Pomorski, A., Schmid, S., Aksimentiev, A., Anslyn, E. V, Bedran, G., Cao, C., Chinappi, M., Coyaud, E., Dekker, C., Dittmar, G., Drachman, N., Eelkema, R., Goodlett, D., Hentz, S., Kalathiya, U., Kelleher, N.L., Kelly, R.T., Kelman, Z., Kim, S.H., Kuster, B., Rodriguez-Larrea, D., Lindsay, S., Maglia, G., Marcotte, E.M., Marino, J.P., Masselon, C., Mayer, M., Samaras, P., Sarthak, K., Sepiashvili, L., Stein, D., Wanunu, M., Wilhelm, M., Yin, P., Meller, A., Joo, C., 2021. The emerging landscape of single-molecule protein sequencing technologies. Nat. Methods 18, 604–617. https://doi.org/10.1038/s41592-021-01143-1

Angel, T.E., Aryal, U.K., Hengel, S.M., Baker, E.S., Kelly, R.T., Robinson, E.W., Smith, R.D., 2012. Mass spectrometry-based proteomics: existing capabilities and future directions. Chem. Soc. Rev. 41, 3912–3928. https://doi.org/10.1039/c2cs15331a

Antonov, A. V, Dietmann, S., Mewes, H.W., 2008. KEGG spider: interpretation of genomics data in the context of the global gene metabolic network. Genome Biol. 9, R179. https://doi.org/10.1186/gb-2008-9-12-r179

Babarinde, I.A., Li, Y., Hutchins, A.P., 2019. Computational Methods for Mapping, Assembly and Quantification for Coding and Non-coding Transcripts. Comput. Struct. Biotechnol. J. 17, 628–637. https://doi.org/https://doi.org/10.1016/j.csbj.2019.04.012

Barrera-Redondo, J., Piñero, D., Eguiarte, L.E., 2020. Genomic, Transcriptomic and Epigenomic Tools to Study the Domestication of Plants and Animals: A Field Guide for Beginners. Front. Genet. 11, 1–24. https://doi.org/10.3389/fgene.2020.00742

Beck, L., Geiger, T., 2022. MS-based technologies for untargeted single-cell proteomics. Curr. Opin. Biotechnol. 76, 102736. https://doi.org/https://doi.org/10.1016/j.copbio.2022.102736

Bhagwat, M., Young, L., Robison, R.R., 2012. Using BLAT to find sequence similarity in closely related genomes. Curr. Protoc. Bioinforma. Chapter 10, 10.8.1-10.8.24. https://doi.org/10.1002/0471250953.bi1008s37

Bumgarner, R., 2013. Overview of DNA microarrays: types, applications, and their future. Curr. Protoc. Mol. Biol. Chapter 22, Unit 22.1. https://doi.org/10.1002/0471142727.mb2201s101

Calderón-González, K.G., Hernández-Monge, J., Herrera-Aguirre, M.E., Luna-Arias, J.P., 2016. Bioinformatics Tools for Proteomics Data Interpretation BT - Modern Proteomics – Sample Preparation, Analysis and Practical Applications, in: Mirzaei, H., Carrasco, M. (Eds.),. Springer International Publishing, Cham, pp. 281–341. https://doi.org/10.1007/978-3-319-41448-5_16

Callahan, N., Tullman, J., Kelman, Z., Marino, J., 2020. Strategies for Development of a Next-Generation Protein Sequencing Platform. Trends Biochem. Sci. 45, 76–89. https://doi.org/10.1016/j.tibs.2019.09.005

Cánovas, F.M., Dumas-Gaudot, E., Recorbet, G., Jorrin, J., Mock, H.-P., Rossignol, M., 2004. Plant proteome analysis. Proteomics 4, 285–298. https://doi.org/10.1002/pmic.200300602

Chen, C., Hou, J., Tanner, J.J., Cheng, J., 2020. Bioinformatics methods for mass spectrometry-based proteomics data analysis. Int. J. Mol. Sci. 21. https://doi.org/10.3390/ijms21082873

Chen, H., Yin, X., Guo, L., Yao, J., Ding, Y., Xu, X., Liu, L., Zhu, Q.-H., Chu, Q., Fan, L., 2021. PlantscRNAdb: A database for plant single-cell RNA analysis. Mol. Plant 14, 855–857. https://doi.org/https://doi.org/10.1016/j.molp.2021.05.002

Chen, W., Yin, X., Mu, J., Yin, Y., 2007. Subfemtomole level protein sequencing by Edman degradation carried out in a microfluidic chip. Chem. Commun. 2488–2490. https://doi.org/10.1039/B700200A

Chu, Q., Zhang, X., Zhu, X., Liu, C., Mao, L., Ye, C., Zhu, Q.-H., Fan, L., 2017. PlantcircBase: A Database for Plant Circular RNAs. Mol. Plant 10, 1126–1128. https://doi.org/https://doi.org/10.1016/j.molp.2017.03.003

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., Szcześniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., Mortazavi, A., 2016. A survey of best practices for RNA-seq data analysis. Genome Biol. 17, 13. https://doi.org/10.1186/s13059-016-0881-8

de Graaf, S.C., Hoek, M., Tamara, S., Heck, A.J.R., 2022. A perspective toward mass spectrometry-based de novo sequencing of endogenous antibodies. MAbs 14, 2079449. https://doi.org/10.1080/19420862.2022.2079449

de Sena Brandine, G., Smith, A.D., 2019. Falco: high-speed FastQC emulation for quality control of sequencing data. F1000Research. https://doi.org/10.12688/f1000research.21142.2

DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.-D., Williams, C., Reich, M., Winckler, W., Getz, G., 2012. RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28, 1530–1532. https://doi.org/10.1093/bioinformatics/bts196

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003a. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4, P3. https://doi.org/10.1186/gb-2003-4-5-p3

Dennis, G., Sherman, B.T., Hosack, D.A., Yang, J., Gao, W., Lane, H.C., Lempicki, R.A., 2003b. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 4, R60. https://doi.org/10.1186/gb-2003-4-9-r60

Deswal, R., Gupta, R., Dogra, V., Singh, R., Abat, J.K., Sarkar, A., Mishra, Y., Rai, V., Sreenivasulu, Y., Amalraj, R.S., Raorane, M., Chaudhary, R.P., Kohli, A., Giri, A.P., Chakraborty, N., Zargar, S.M., Agrawal, V.P., Agrawal, G.K., Job, D., Renaut, J., Rakwal, R., 2013. Plant proteomics in India and Nepal: current status and challenges ahead. Physiol. Mol. Biol. Plants 19, 461–477. https://doi.org/10.1007/s12298-013-0198-y

Dong, H., Zhang, A., Sun, H., Wang, H., Lu, X., Wang, M., Ni, B., Wang, X., 2012. Ingenuity pathways analysis of urine metabolomics phenotypes toxicity of Chuanwu in Wistar rats by UPLC-Q-TOF-HDMS coupled with pattern recognition methods. Mol. Biosyst. 8, 1206–1221. https://doi.org/10.1039/C1MB05366C

Dupree, E.J., Jayathirtha, M., Yorkey, H., Mihasan, M., Petre, B.A., Darie, C.C., 2020. A critical review of bottom-up proteomics: The good, the bad, and the future of this field. Proteomes 8, 1–26. https://doi.org/10.3390/proteomes8030014

Feng, J., Meyer, C.A., Wang, Q., Liu, J.S., Shirley Liu, X., Zhang, Y., 2012. GFOLD: a generalized fold change for ranking differentially expressed genes from RNA-seq data. Bioinformatics 28, 2782–

2788. https://doi.org/10.1093/bioinformatics/bts515
Geniza, M., Jaiswal, P., 2017. Tools for building de novo transcriptome assembly. Curr. Plant Biol. 11–12, 41–45. https://doi.org/https://doi.org/10.1016/j.cpb.2017.12.004
Ghosh, S., Chan, C.-K.K., 2016. Analysis of RNA-Seq Data Using TopHat and Cufflinks. Methods Mol. Biol. 1374, 339–361. https://doi.org/10.1007/978-1-4939-3167-5_18
Giacomello, S., 2021. A new era for plant science: spatial single-cell transcriptomics. Curr. Opin. Plant Biol. 60, 102041. https://doi.org/https://doi.org/10.1016/j.pbi.2021.102041
Guo, J., Huang, Z., Sun, J., Cui, X., Liu, Y., 2021. Research Progress and Future Development Trends in Medicinal Plant Transcriptomics. Front. Plant Sci. 12, 691838. https://doi.org/10.3389/fpls.2021.691838
Hale, J.E., 2013. Advantageous Uses of Mass Spectrometry for the Quantification of Proteins. Int. J. Proteomics 2013, 219452. https://doi.org/10.1155/2013/219452
Han, X., Aslanian, A., Yates, J.R. 3rd, 2008. Mass spectrometry for proteomics. Curr. Opin. Chem. Biol. 12, 483–490. https://doi.org/10.1016/j.cbpa.2008.07.024
Hina, F., Yisilam, G., Wang, S., Li, P., Fu, C., 2020. De novo Transcriptome Assembly, Gene Annotation and SSR Marker Development in the Moon Seed Genus Menispermum (Menispermaceae). Front. Genet. 11, 1–13. https://doi.org/10.3389/fgene.2020.00380
Hou, Y.P., Diao, T.T., Xu, Z.H., Mao, X.Y., Wang, C., Li, B., 2022. Bioinformatic Analysis Combined With Experimental Validation Reveals Novel Hub Genes and Pathways Associated With Focal Segmental Glomerulosclerosis. Front. Mol. Biosci. 8, 1–9. https://doi.org/10.3389/fmolb.2021.691966
Hu, J., Rampitsch, C., Bykova, N. V., 2015. Advances in plant proteomics toward improvement of crop productivity and stress resistance. Front. Plant Sci. 6, 1–15. https://doi.org/10.3389/fpls.2015.00209
Hunt, D.F., Yates, J.R. 3rd, Shabanowitz, J., Winston, S., Hauer, C.R., 1986. Protein sequencing by tandem mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. 83, 6233–6237. https://doi.org/10.1073/pnas.83.17.6233
Jammali, S., Aguilar, J.-D., Kuitche, E., Ouangraoua, A., 2019. SplicedFamAlign: CDS-to-gene spliced alignment and identification of transcript orthology groups. BMC Bioinformatics 20, 133. https://doi.org/10.1186/s12859-019-2647-2
Kanehisa, M., Goto, S., 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 28, 27–30. https://doi.org/10.1093/nar/28.1.27
Kapustin, Y., Souvorov, A., Tatusova, T., Lipman, D., 2008. Splign: algorithms for computing spliced alignments with identification of paralogs. Biol. Direct 3, 20. https://doi.org/10.1186/1745-6150-3-20
Karagiannis, K., Simonyan, V., Mazumder, R., 2013. SNVDis: A Proteome-wide Analysis Service for Evaluating nsSNVs in Protein Functional Sites and Pathways. Genomics. Proteomics Bioinformatics 11, 122–126. https://doi.org/https://doi.org/10.1016/j.gpb.2012.10.003
Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., Salzberg, S.L., 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 14, R36. https://doi.org/10.1186/gb-2013-14-4-r36
Kosová, K., Vítámvás, P., Urban, M.O., Prášil, I.T., Renaut, J., 2018. Plant abiotic stress proteomics: The major factors determining alterations in cellular proteome. Front. Plant Sci. 9, 1–22. https://doi.org/10.3389/fpls.2018.00122
Kumar, G., Ertel, A., Feldman, G., Kupper, J., Fortina, P., 2020. iSeqQC: a tool for expression-based quality control in RNA sequencing. BMC Bioinformatics 21, 56. https://doi.org/10.1186/s12859-020-3399-8
Kumari, S., Verma, L.K., Weller, J.W., 2007. AffyMAPSDetector: a software tool to characterize Affymetrix GeneChip™ expression arrays with respect to SNPs. BMC Bioinformatics 8, 276. https://doi.org/10.1186/1471-2105-8-276
Leggett, R.M., Ramirez-Gonzalez, R.H., Clavijo, B.J., Waite, D., Davey, R.P., 2013. Sequencing quality assessment tools to enable data-driven informatics for high throughput genomics. Front. Genet. 4, 288. https://doi.org/10.3389/fgene.2013.00288
Li, Y., Ge, X., Peng, F., Li, W., Li, J.J., 2022. Exaggerated false positives by popular differential expression methods when analyzing human population samples. Genome Biol. 23, 79. https://doi.org/10.1186/s13059-022-02648-4
Libault, M., Pingault, L., Zogli, P., Schiefelbein, J., 2017. Plant Systems Biology at the Single-Cell Level. Trends Plant Sci. 22, 949–960. https://doi.org/https://doi.org/10.1016/j.tplants.2017.08.006
Liu, Y., Lu, S., Liu, K., Wang, S., Huang, L., Guo, L., 2019. Proteomics: a powerful tool to study plant responses to biotic stress. Plant Methods 15, 135. https://doi.org/10.1186/s13007-019-0515-8

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., Shafee, T., 2017. Transcriptomics technologies. PLOS Comput. Biol. 13, e1005457. https://doi.org/10.1371/journal.pcbi.1005457

Macklin, A., Khan, S., Kislinger, T., 2020. Recent advances in mass spectrometry based clinical proteomics: applications to cancer research. Clin. Proteomics 17, 17. https://doi.org/10.1186/s12014-020-09283-w

Mahmood, K., Orabi, J., Kristensen, P.S., Sarup, P., Jørgensen, L.N., Jahoor, A., 2020. De novo transcriptome assembly, functional annotation, and expression profiling of rye (*Secale cereale* L.) hybrids inoculated with ergot (*Claviceps purpurea*). Sci. Rep. 10, 13475. https://doi.org/10.1038/s41598-020-70406-2

Malone, J.H., Oliver, B., 2011. Microarrays, deep sequencing and the true measure of the transcriptome. BMC Biol. 9, 34. https://doi.org/10.1186/1741-7007-9-34

Martin, L.B.B., Fei, Z., Giovannoni, J.J., Rose, J.K.C., 2013. Catalyzing plant science research with RNA-seq. Front. Plant Sci. 4, 1–10. https://doi.org/10.3389/fpls.2013.00066

Mergner, J., Kuster, B., 2022. Plant Proteome Dynamics. Annu. Rev. Plant Biol. 73, 67–92. https://doi.org/10.1146/annurev-arplant-102620-031308

Mi, H., Muruganujan, A., Casagrande, J.T., Thomas, P.D., 2013. Large-scale gene function analysis with the PANTHER classification system. Nat. Protoc. 8, 1551–1566. https://doi.org/10.1038/nprot.2013.092

Miyashita, M., Presley, J.M., Buchholz, B.A., Lam, K.S., Lee, Y.M., Vogel, J.S., Hammock, B.D., 2001. Attomole level protein sequencing by Edman degradation coupled with accelerator mass spectrometry. Proc. Natl. Acad. Sci. U. S. A. 98, 4403–4408. https://doi.org/10.1073/pnas.071047998

Moreno-Santillán, D.D., Machain-Williams, C., Hernández-Montes, G., Ortega, J., 2019. De Novo Transcriptome Assembly and Functional Annotation in Five Species of Bats. Sci. Rep. 9, 6222. https://doi.org/10.1038/s41598-019-42560-9

Pandeswari, P.B., Sabareesh, V., 2019. Middle-down approach: a choice to sequence and characterize proteins/proteomes by mass spectrometry. RSC Adv. 9, 313–344. https://doi.org/10.1039/C8RA07200K

Patole, C., Bindschedler, L. V, 2019. Chapter 4 - Plant proteomics: A guide to improve the proteome coverage, in: Meena, S.N., Naik, M.M.B.T.-A. in B.S.R. (Eds.),. Academic Press, pp. 45–67. https://doi.org/https://doi.org/10.1016/B978-0-12-817497-5.00004-5

Pierlé, S.A., Dark, M.J., Dahmen, D., Palmer, G.H., Brayton, K.A., 2012. Comparative genomics and transcriptomics of trait-gene association. BMC Genomics 13, 669. https://doi.org/10.1186/1471-2164-13-669

Raghavan, V., Kraft, L., Mesny, F., Rigerte, L., 2022. A simple guide to de novo transcriptome assembly and annotation. Brief. Bioinform. 23. https://doi.org/10.1093/bib/bbab563

Ragoussis, J., Elvidge, G., 2006. Affymetrix GeneChip® system: moving from research to the clinic. Expert Rev. Mol. Diagn. 6, 145–152. https://doi.org/10.1586/14737159.6.2.145

Rajczewski, A.T., Jagtap, P.D., Griffin, T.J., 2022. An overview of technologies for MS-based proteomics-centric multi-omics. Expert Rev. Proteomics 19, 165–181. https://doi.org/10.1080/14789450.2022.2070476

Rao, M.S., Van Vleet, T.R., Ciurlionis, R., Buck, W.R., Mittelstadt, S.W., Blomme, E.A.G., Liguori, M.J., 2019. Comparison of RNA-Seq and microarray gene expression platforms for the toxicogenomic evaluation of liver from short-term rat toxicity studies. Front. Genet. 10, 1–16. https://doi.org/10.3389/fgene.2018.00636

Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140. https://doi.org/10.1093/bioinformatics/btp616

Rodriques, S.G., Marblestone, A.H., Boyden, E.S., 2019. A theoretical analysis of single molecule protein sequencing via weak binding spectra. PLoS One 14, e0212868. https://doi.org/10.1371/journal.pone.0212868

Sessegolo, C., Cruaud, C., Da Silva, C., Cologne, A., Dubarry, M., Derrien, T., Lacroix, V., Aury, J.-M., 2019. Transcriptome profiling of mouse samples using nanopore sequencing of cDNA and RNA molecules. Sci. Rep. 9, 14908. https://doi.org/10.1038/s41598-019-51470-9

Simpson, J.T., Wong, K., Jackman, S.D., Schein, J.E., Jones, S.J.M., Birol, I., 2009. ABySS: a parallel assembler for short read sequence data. Genome Res. 19, 1117–1123. https://doi.org/10.1101/gr.089532.108

Singhal, N., Kumar, M., Kanaujia, P.K., Virdi, J.S., 2015. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. Front. Microbiol. 6, 1–16. https://doi.org/10.3389/fmicb.2015.00791

Smythers, A.L., Hicks, L.M., 2021. Mapping the plant proteome: tools for surveying coordinating pathways. Emerg. Top. Life Sci. 5, 203–220. https://doi.org/10.1042/ETLS20200270

Spies, D., Ciaudo, C., 2015. Dynamics in Transcriptomics: Advancements in RNA-seq Time Course and Downstream Analysis. Comput. Struct. Biotechnol. J. 13, 469–477. https://doi.org/https://doi.org/10.1016/j.csbj.2015.08.004

Squair, J.W., Gautier, M., Kathe, C., Anderson, M.A., James, N.D., Hutson, T.H., Hudelle, R., Qaiser, T., Matson, K.J.E., Barraud, Q., Levine, A.J., La Manno, G., Skinnider, M.A., Courtine, G., 2021. Confronting false discoveries in single-cell differential expression. Nat. Commun. 12, 5692. https://doi.org/10.1038/s41467-021-25960-2

Standing, K.G., 2003. Peptide and protein de novo sequencing by mass spectrometry. Curr. Opin. Struct. Biol. 13, 595–601. https://doi.org/https://doi.org/10.1016/j.sbi.2003.09.005

Suriyakala, G., Sathiyaraj, S., Gandhi, A.D., Vadakkan, K., Mahadeva Rao, U.S., Babujanarthanam, R., 2021. Plumeria pudica Jacq. flower extract - mediated silver nanoparticles: Characterization and evaluation of biomedical applications. Inorg. Chem. Commun. 126, 108470. https://doi.org/10.1016/j.inoche.2021.108470

Tamara, S., den Boer, M.A., Heck, A.J.R., 2022. High-Resolution Native Mass Spectrometry. Chem. Rev. 122, 7269–7326. https://doi.org/10.1021/acs.chemrev.1c00212

Trapnell, C., Pachter, L., Salzberg, S.L., 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25, 1105–1111. https://doi.org/10.1093/bioinformatics/btp120

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., Pachter, L., 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. 28, 511–515. https://doi.org/10.1038/nbt.1621

Ungaro, A., Pech, N., Martin, J.-F., McCairns, R.J.S., Mévy, J.-P., Chappaz, R., Gilles, A., 2017. Challenges and advances for transcriptome assembly in non-model species. PLoS One 12, e0185020. https://doi.org/10.1371/journal.pone.0185020

Vadakkan, K., 2020. Molecular Mechanism of Bacterial Quorum Sensing and Its Inhibition by Target Specific Approaches. ACS Symp. Ser. 1374, 21–234. https://doi.org/10.1021/bk-2020-1374.ch012

Vadakkan, K., 2019. Acute and sub-acute toxicity study of bacterial signaling inhibitor *Solanum torvum* root extract in Wister rats. Clin. Phytoscience 5. https://doi.org/10.1186/s40816-019-0113-3

Vadakkan, K., Cheruvathur, M.K., Chulliparambil, A.S., Francis, F., Abimannue, A.P., 2021. Proteolytic enzyme arbitrated antagonization of helminthiasis by *Cinnamomum cappara* leaf extract in Pheretima posthuma. Clin. Phytoscience 7. https://doi.org/10.1186/s40816-021-00261-9

Vadakkan, Kayeen, Choudhury, A.A., Gunasekaran, R., Hemapriya, J., Vijayanand, S., 2018a. Quorum sensing intervened bacterial signaling: Pursuit of its cognizance and repression. J. Genet. Eng. Biotechnol. 16, 239–252. https://doi.org/10.1016/j.jgeb.2018.07.001

Vadakkan, Kayeen, Choudhury, A.A., Gunasekaran, R., Hemapriya, J., Vijayanand, S., 2018b. Quorum sensing intervened bacterial signaling: Pursuit of its cognizance and repression. J. Genet. Eng. Biotechnol. 16, 239–252. https://doi.org/10.1016/j.jgeb.2018.07.001

Vadakkan, K., Gunasekaran, R., Choudhury, A.A., Ravi, A., Arumugham, S., Hemapriya, J., Vijayanand, S., 2018. Response Surface Modelling through Box-Behnken approach to optimize bacterial quorum sensing inhibitory action of Tribulus terrestris root extract. Rhizosphere 6, 134–140. https://doi.org/10.1016/j.rhisph.2018.06.005

Vadakkan, K., Hemapriya, J., Anbarasu, A., Ramaiah, S., Vijayanand, S., 2020. Quorum quenching by 2-Hydroxyanisole extracted from *Solanum torvum* on *Pseudomonas aeruginosa* and its inhibitory action upon LasR protein. Gene Reports 21, 100802. https://doi.org/10.1016/j.genrep.2020.100802

Vadakkan, K., Hemapriya, J., Selvaraj, V., 2019a. Quorum quenching intervened in vivo attenuation and immunological clearance enhancement by *Solanum torvum* root extract against *Pseudomonas aeruginosa* instigated pneumonia in Sprague Dawley rats. Clin. Phytoscience 5, 24. https://doi.org/10.1186/s40816-019-0120-4

Vadakkan, K., Hemapriya, J., Selvaraj, V., 2019b. Quorum quenching intervened in vivo attenuation and immunological clearance enhancement by *Solanum torvum* root extract against *Pseudomonas aeruginosa* instigated pneumonia in Sprague Dawley rats.

Vadakkan, Kayeen, Vijayanand, S., Choudhury, A.A., Gunasekaran, R., Hemapriya, J., 2018c. Optimization of quorum quenching mediated bacterial attenuation of *Solanum torvum* root

extract by response surface modelling through Box-Behnken approach. J. Genet. Eng. Biotechnol. https://doi.org/10.1016/j.jgeb.2018.02.001

Vadakkan, Kayeen, Vijayanand, S., Choudhury, A.A., Gunasekaran, R., Hemapriya, J., 2018d. Optimization of quorum quenching mediated bacterial attenuation of *Solanum torvum* root extract by response surface modelling through Box-Behnken approach. J. Genet. Eng. Biotechnol. 16, 381–386. https://doi.org/10.1016/j.jgeb.2018.02.001

Vadakkan, K., Vijayanand, S., Hemapriya, J., Gunasekaran, R., 2019c. Quorum sensing inimical activity of Tribulus terrestris against gram negative bacterial pathogens by signalling interference. 3 Biotech 9, 163. https://doi.org/10.1007/s13205-019-1695-7

Vecchi, M.M., Xiao, Y., Wen, D., 2019. Identification and Sequencing of N-Terminal Peptides in Proteins by LC-Fluorescence-MS/MS: An Approach to Replacement of the Edman Degradation. Anal. Chem. 91, 13591–13600. https://doi.org/10.1021/acs.analchem.9b02754

Vitorino, R., Guedes, S., Trindade, F., Correia, I., Moura, G., Carvalho, P., Santos, M.A.S., Amado, F., 2020. De novo sequencing of proteins by mass spectrometry. Expert Rev. Proteomics 17, 595–607. https://doi.org/10.1080/14789450.2020.1831387

Walker, J.M., 1997. The Dansyl-Edman Method for Manual Peptide Sequencing BT - Protein Sequencing Protocols, in: Smith, B.J. (Ed.),. Humana Press, Totowa, NJ, pp. 183–187. https://doi.org/10.1385/0-89603-353-8:183

Wang, B., Kumar, V., Olson, A., Ware, D., 2019. Reviving the transcriptome studies: An insight into the emergence of single-molecule transcriptome sequencing. Front. Genet. 10, 1–11. https://doi.org/10.3389/fgene.2019.00384

Wang, Z., Gerstein, M., Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. 10, 57–63. https://doi.org/10.1038/nrg2484

Winck, F.V., dos Santos, A.L.W., Calderan-Rodrigues, M.J., 2021. Plant Proteomics and Systems Biology BT - Advances in Plant Omics and Systems Biology Approaches, in: Vischi Winck, F. (Ed.),. Springer International Publishing, Cham, pp. 51–66. https://doi.org/10.1007/978-3-030-80352-0_3

Wu, T.D., Watanabe, C.K., 2005. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. Bioinformatics 21, 1859–1875. https://doi.org/10.1093/bioinformatics/bti310

Xie, H., Wang, W., Sun, F., Deng, K., Lu, X., Liu, H., Zhao, W., Zhang, Y., Zhou, X., Li, K., Hou, Y., 2017. Proteomics analysis to reveal biological pathways and predictive proteins in the survival of high-grade serous ovarian cancer. Sci. Rep. 7, 9896. https://doi.org/10.1038/s41598-017-10559-9

Yang, I.S., Kim, S., 2015. Analysis of Whole Transcriptome Sequencing Data: Workflow and Software. Genomics Inform. 13, 119–125. https://doi.org/10.5808/GI.2015.13.4.119

Yang, M., Wang, Q., Wang, S., Wang, Y., Zeng, Q., Qin, Q., 2019. Transcriptomics analysis reveals candidate genes and pathways for susceptibility or resistance to Singapore grouper iridovirus in orange-spotted grouper (*Epinephelus coioides*). Dev. Comp. Immunol. 90, 70–79. https://doi.org/10.1016/j.dci.2018.09.003

Yang, Y., Smith, S.A., 2013. Optimizing de novo assembly of short-read RNA-seq data for phylogenomics. BMC Genomics 14, 328. https://doi.org/10.1186/1471-2164-14-328

Yu, J., Gu, X., Yi, S., 2016. Ingenuity pathway analysis of gene expression profiles in distal nerve stump following nerve injury: Insights into wallerian degeneration. Front. Cell. Neurosci. 10, 1–12. https://doi.org/10.3389/fncel.2016.00274

Zhou, Q., Su, X., Jing, G., Chen, S., Ning, K., 2018. RNA-QC-chain: comprehensive and fast quality control for RNA-Seq data. BMC Genomics 19, 144. https://doi.org/10.1186/s12864-018-4503-6

How to cite this article:

Nimitha, P. A. 2024. An OMICS-Based Approach Studies Natural Products. *Int.J.Curr.Microbiol.App.Sci.* 13(01): 94-103. **doi:** https://doi.org/10.20546/ijcmas.2024.1301.012