

Advancement in Medicinal plants transcriptomic approaches for deciphering bioactive secondary metabolites

Awais ALi*, Umar Khitab Khan, Waseef Ullah, Muneeba Aslam

Abstract—Plants have been used for medical purposes for centuries, having cures and treatments for many diseases. They developed in the 19th century, which slowed most of the problems of the modern era. Transcriptome treats diseases to improve crop yield to produce better-designed crops in the breeding program. The Transcriptome classifies organisms based on their transcriptomic approaches, which leads to the development of drugs and medicine. The review aims to elaborate on all the approaches and gaps in plant transcriptome and NGS to retrieve all possible gaps and medicinal value of the transcriptome approaches in *Arabidopsis*, *Salvia miltiorrhiza*, and *Artemisia*. In addition We will focus on transcriptome technologies and bioinformatics tools to approach medicinal plants, which are used for medical purposes like anti-inflammatory, anticancer, anti-aging, etc., and disease resistance compounds and to improve fruit quality (i.e. fruit shape, sweetness, flavor, and size) To be concluded This study will point out the potential of genetic, evaluation of transcriptome data to enhance the medicinal compounds discovery in new organisms.

Index Terms transcriptome, drugs, NGS, genetics, bioinformatics, medicinal plants, breeding, disease

1. INTRODUCTION:

Medicinal and aromatic plants are well known for producing secondary metabolites which are used for different purposes like dyes, fragrances, drugs, dyes, fragrances, and insecticides. The transcriptome analysis is done in order to explore the Denovo pathways of secondary metabolites in plants. [1] These secondary metabolites are important therapeutic characteristics that are common in medicinal plants. Until recently, efforts to unravel biosynthetic pathways were limited by a lack of plant genomics resources. Recent advances in the generation of high-throughput "omics" datasets, computational tools, and, in addition, more analytical methodologies have resulted in the explanation of massive amounts of plant bioactive metabolites. Through synthetic biology, researchers have taken a step forward in establishing a more sustainable supply of these biomolecules. [2]. Recent advancements in next-generation sequencing (NGS) technology have boosted medicinal plant research. NGS methods enable high-throughput whole-genome sequencing and direct RNA sequencing. The produced sequence data sets can further be studied for use in other fields of study such as data mining for short and long non-coding RNAs, mining of molecular markers, gene identification, and so on. [3] A comprehensive transcriptome analysis for medicinal plants created using next generation sequencing. Mining various route analysis could help with pathway elaboration and gene/enzyme discovery. The utility of transcriptome techniques in identifying the pathways involved in the synthesis of important secondary metabolites as (Show in Table 1) in medicinal plants is also investigated [4]. *Salvia miltiorrhiza*, which contains tanshinones and phenolic acids as bioactive components, which long been used to treat cardiovascular and cerebrovascular disorders. Because of the effective genetic transformation method, simplified hairy roots system, and high-throughput sequencing, biosynthetic study of secondary metabolites in *S. miltiorrhiza* has made significant progress. Metabolic engineeri

ng or synthetic biology will be able to enhance or manufacture substances in plants or microorganisms after analyzing the biosynthetic process [5]. *Staphylococcus aureus* is a common human pathogen and commensal, with the human nose serving as the primary reservoir. The metatranscriptome examined using RNA-seq after RNA was collected from the anterior nares of three verified *S. aureus* carriers. Survival in the nose heavily influenced by iron restriction, as seen by the expression of iron acquisition mechanisms. The assessment of general in vivo expression patterns and similarities between various strains will result in new knowledge-based techniques for colonisation management in the future [6]. *Artemisia absinthium* essential oil (AEO) shows the direct and indirect plant protection effects of an AEO on tomato seedlings against *Fusarium oxysporum* sp. *oxysporum radialis lycopersici* (Fol) and its treatment could induce de novo epigenetic changes in tomato, modulating the speed and extent of its immune response to Fol. The EO-seed coating could be a new strategy to prime durable tomato resistance, compatible with other environmentally friendly bio pesticides. Transcriptome Assembly and Function Annotation Methods [7]

Table 1: secondary metabolites found in medicinal plants from RNA-seq analysis.

S. NO	Metabolite	Medicinal properties	Plant	Class	Reference
01	Neral	Treatment of nervous disorders	<i>lemon myrtle</i>	Monoterpe noid	[8]
02	Agrimoniin	antiparasitic , antidiarrheal anticancer	<i>Agrimonia pilosa</i>	Tannin	[9]
03	Rutin		<i>Egyptian flora</i>		[10]
04	Ellagitannin	Anti-inflammatory	<i>Agrimonia pilosa</i>	Tannin	[9]
05	Geranial	Anti-inflammatory	<i>lemon myrtle</i>	Monoterpe noid	[8]
06	Resveratrol	Anticancer	<i>Egyptian flora</i>		[10]
07	Pelargonidin	Anti-inflammatory, Anti-cancer	<i>berry fruits</i>	Flavonoids	[11]
08	Pelargonidin-3-O-glucoside	Anti-cancer	<i>berry fruits</i>	Flavonoids	[12]
09	Beta-Amyrin	Anti-inflammatory and Antioxidant		Terpenoid	[13]
10	Lupeol	anti-inflammatory, anti-cancer		Terpenoid	[14]
11	Galangin	Ant mutagenic, Antioxidant,		Flavonoids	[15]
12	Tanshinones	Alleviate infarct size, reduce edema	<i>Salvia miltiorrhiza</i>	Diterpene	[5]
13	Alliin	Antifungal, Antiprotozoal, Antioxidant,	<i>Allium sativum L</i>	Alkaloid	[16]

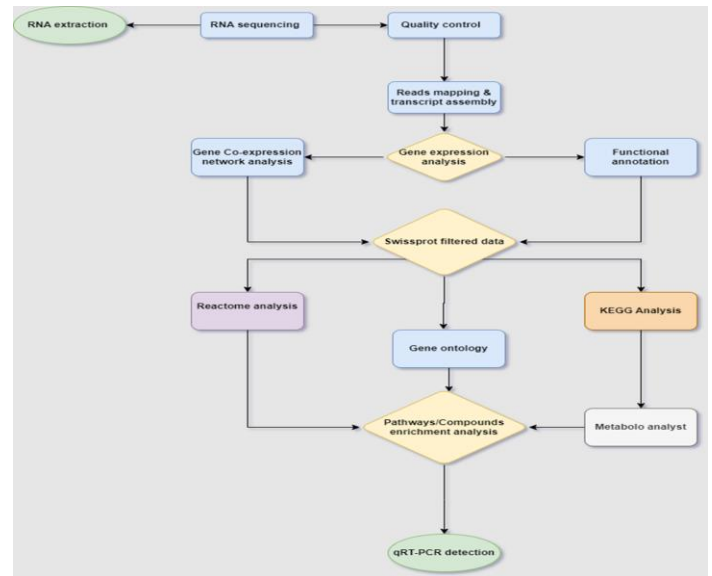


Figure 2: the flow chart of transcriptomic data processing from RNA extraction to and pathway enrichment and compounds biosynthesis.

The studies show the transcriptomic of tomato seeds coated with an antifungal EO against the phytopathogen *Fusarium oxysporum*. The interaction of tomato with *Fusarium oxysporum* f. sp. *lycopersici* (Fol) used as a model to investigate the molecular basis of disease resistance and susceptibility. [17], [18] Here, we discuss novel metabolites and molecular alterations linked to tomato plants' short- and long-term immunity to plants germinated from AEO-coated against the very damaging phytopathogenic fungus Fol. Based on metabolomics and RNA-seq analyses, the consequences of the EO described. This study suggests new applications for EOs as a source of bio-pesticides and biotechnological tools that are safe for the environment.

1. Transcriptomic analysis:

1. Plant materials

For the extraction of the RNA-seq experiment, *Arabidopsis thaliana* had used for the experiment. In this process, many seeds have grown in a growth chamber with a time-lapse of 1/10 hours' light and dark cycle at 23°C. The seeds were grown for three weeks, then removed from the soil, rinsed, and submerged at 1000pm with gentle shaking for 16 hours A Mock solution were used as control. For static analysis, all experiments were duplicated and the RNA was extracted.[19]

1.1. RNA extraction, library construction and sequencing:

Transcriptome assembly is required for a variety of subsequent analyses. Raw transcriptome data is error-prone due to the large amount of transcriptome data; as a result, selecting assembly software based on different transcriptomic data and research objectives is critical. The first consideration is whether a reference sequence is available; assembly is classified into two types based on this: de novo assembly and reference sequence-based assembly. Because most medicinal plants are non-model organisms, genome sequence information is frequently lacking; thus, reference assembly cannot be performed. The only assembly method suitable for non-model plants is de novo assem-

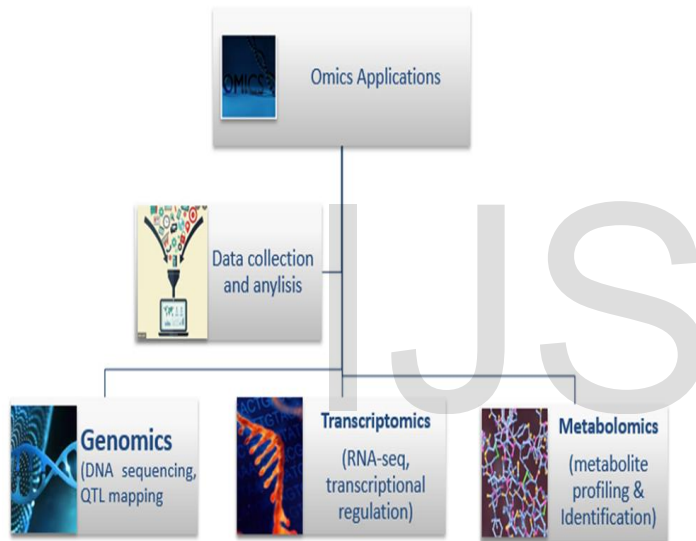


Figure 1: different fields of OMICS and application for computational data analysis.

bly.[20].

1.2. Post-sequence analysis:

After high capacity sequencing, reads with ambiguous bases, short, prematurely ended sequences, and bad quality reads eliminated from the raw sequence reads. Using TopHat (v1.3.1) BWA software, clean reads aligned to the Arabidopsis genome assembly TAIR10 (used as a reference). [21] Alignment results were stored in the SAM/BAM file format (Nakamura et al., 2016) to allow for downstream analysis Cufflinks (v1.0.3) software was then used to identify transcripts and their abundance levels in RNA-seq data, as well as to perform differential gene expression analysis across many samples[22]. Cufflinks calculate expression levels in fragments per kilo base of sequence per million fragments mapped (FPKM), which were used to compare treatment effects on gene expression.[22]

1.3. Primary and secondary metabolite analysis:

Gene Ontology (GO) enrichment performed by using Panther GO3. The GO enrichment analysis provided all the GO terms that significantly enriched in the DEGs relative to the genomic background, and the DEGs were filtering according to cellular components, molecular functions, and biological processes. The Kyoto Encyclopedia of Genes and Genomes (KEGG) is a main pathway-related database. Based on the comparison of the DEGs to the genomic background, pathway enrichment analysis pinpointed the enriched pathways available at (<http://www.genome.jp/kegg/>). (Li et al., 2019). KEGG used to retrieve all possible pathways of secondary metabolites and other pathways. As shown in Table 2; KEGG shows pathway enrichment with uploaded data from any organism [23], [24]. MetaCyc/ PlantCyc (version 3.0) includes 714 metabolic pathways and 2,619 reactions from over 300 species. PlantCyc(<http://plantcyc.org>). and MetaCyc provide a full Enrichment analysis of metabolic pathway database.[25]. MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca>) implements a knowledge-based network approach, whereby user's data (metabolites and genes) can be projected onto five existing biological networks: 1) Pathway-based network discovery, 2) Gene-metabolite interaction network, 3) Metabolite-phenotype interaction network, 4) Metabolite-metabolite interaction network, and 5) Metabolite-gene-phenotype interaction network.[26]. The Plant Reactome is an open-source, comparative plant pathway knowledgebase of the Gramene project (<https://plantreactome.org>). For manual pathway curation, *Oryza sativa* is used as a reference species, and the Reactome data model and framework are used to project gene-orthology onto another 82 plant species to expand route knowledge. Users can upload and analyse their own omics data, such as data on gene expression, and then overlay curated or experimental data on gene-gene interactions to increase their understanding of circuit structure.[27].

Table 2: show different databases uses for different analysis with their distinctive features

Database	Biosynthesis of metabolites	Enrichment analysis	Network interaction	Gene expression in different parts
KEGG	✓	✓	✗	✓
MetaboAnalyst5.0	✗	✗	✓	✓
GO/GO3	✗	✓	✗	✗
Metacyc	✓	✓	✗	✗
Panther	✓	✓	✗	✗
Plant reactome	✓	✓	✗	✓

1.4. qRT-PCR Detection:

The stem-loop qRT-PCR approach presented here is intended to detect and quantify mature miRNAs quickly, specifically, accurately, and consistently. Just before reverse transcription, a miRNA-specific stem-loop RT primer is hybridised to the miRNA. Using a forward primer specific to miRNA and a universal reverse primer, the RT product is then amplified and analysed in real time. With this technique, miRNA expression profiles can be determined with as little as 10 pg of total RNA and at a relatively high throughput.

2. CONCLUSION

The study reveals methods and applications of plant transcriptome analysis to generate high-quality RNA-Seq data, to explore the biosynthetic pathways of secondary metabolites including Terpenoid, and alkaloids from unknown plants, and to increase the field yield of different medicinal compounds. Different databases perform different analysis revealed pathways enriched which were mainly involved in biosynthetic pathways of secondary metabolites from RNA-seq Data.

3. REFERENCES

- [1] S. G. Gandhi, V. Mahajan, and Y. S. Bedi, "Changing trends in biotechnology of secondary metabolism in medicinal and aromatic plants," *Planta*, vol. 241, no. 2, pp. 303–317, 2015.
- [2] S. Pathak, A. V. Agarwal, P. Agarwal, and P. K. Trivedi, "Secondary Metabolite Pathways in Medicinal Plants: Approaches in Reconstruction and Analysis," in *Molecular Approaches in Plant Biology and Environmental Challenges*, Springer, 2019, pp. 339–364.
- [3] S. Chaudhary and P. C. Sharma, "Next generation sequencing-based exploration of genomes and transcriptomes of medicinal plants," *Indian J. Plant Physiol.*, vol. 21, no. 4, pp. 489–503, 2016.
- [4] S. Tripathi, J. S. Jadaun, M. Chandra, and N. S. Sangwan, "Medicinal plant transcriptomes: the new gateways for accelerated understanding of plant secondary metabolism," *Plant Genet. Resour.*, vol. 14, no. 4, pp. 256–269, 2016.
- [5] X.-H. Ma et al., "The biosynthetic pathways of tanshinones and phenolic acids in *Salvia miltiorrhiza*," *Molecules*, vol. 20, no. 9, pp. 16235–16254, 2015.
- [6] D. Chaves-Moreno, M. L. Wos-Oxley, R. Jáuregui, E. Medina, A. Oxley, and D. H. Pieper, "Exploring the transcriptome of *Staphylococcus aureus* in its natural niche," *Sci. Rep.*, vol. 6, no. 1, pp. 1–11, 2016.
- [7] S. Soudani, C. Poza-Carrión, N. I. C. G. De, A. González-Coloma, M. F. Andrés, and M. Berrocal-Lobo, "Essential Oils Prime Epigenetic and Metabolomic Changes in Tomato Defense Against *Fusarium oxysporum*," *Front. Plant Sci.*, vol. 13, p. 804104, 2022.
- [8] T. Usuki and M. Yoshizawa-Fujita, "Extraction and Isolation of Natural Organic Compounds from Plant Leaves Using Ionic Liquids.," *Adv. Biochem. Eng. Biotechnol.*, vol. 168, pp. 227–240, 2019, doi: 10.1007/10_2018_80.
- [9] D. M. Grochowski et al., "A comprehensive review of agrimoniin.," *Ann. N. Y. Acad. Sci.*, vol. 1401, no. 1, pp. 166–180, Aug. 2017, doi: 10.1111/nyas.13421.

- [10] H. R. El-Seedi *et al.*, "Exploring natural products-based cancer therapeutics derived from Egyptian flora," *J. Ethnopharmacol.*, vol. 269, p. 113626, Apr. 2021, doi: 10.1016/j.jep.2020.113626.
- [11] S. Acquah, J. N. Boampong, and B. A. Eghan Jr, "Increased oxidative stress and inflammation independent of body adiposity in diabetic and nondiabetic controls in falciparum malaria," *Biomed Res. Int.*, vol. 2016, 2016.
- [12] A. M. Amini, K. Muzs, J. P. E. Spencer, and P. Yaqoob, "Pelargonidin-3-O-glucoside and its metabolites have modest anti-inflammatory effects in human whole blood cultures," *Nutr. Res.*, vol. 46, pp. 88–95, 2017.
- [13] S. A. Holanda Pinto, L. M. S. Pinto, G. M. A. Cunha, M. H. Chaves, F. A. Santos, and V. S. Rao, "Anti-inflammatory effect of α , β -Amyrin, a pentacyclic triterpene from *Protium heptaphyllum* in rat model of acute periodontitis," *Inflammopharmacology*, vol. 16, no. 1, pp. 48–52, 2008, doi: 10.1007/s10787-007-1609-x.
- [14] M. Saleem, "Lupeol, a novel anti-inflammatory and anti-cancer dietary triterpene," *Cancer Lett.*, vol. 285, no. 2, pp. 109–115, Nov. 2009, doi: 10.1016/j.canlet.2009.04.033.
- [15] M. Bacanlı, A. A. Başaran, and N. Başaran, "Galangin as a plant phenolic and usage in health and disease," in *Polyphenols: prevention and treatment of human disease*, Elsevier, 2018, pp. 433–438.
- [16] G. El-Saber Batiha *et al.*, "Chemical constituents and pharmacological activities of garlic (*Allium sativum* L.): A review," *Nutrients*, vol. 12, no. 3, p. 872, 2020.
- [17] C. E. Windels, "Economic and social impacts of Fusarium head blight: changing farms and rural communities in the Northern Great Plains," *Phytopathology*, vol. 90, no. 1, pp. 17–21, 2000.
- [18] A. O. Carvalho *et al.*, "Cloning and characterization of a cowpea seed lipid transfer protein cDNA: expression analysis during seed development and under fungal and cold stresses in seedlings' tissues," *Plant Physiol. Biochem.*, vol. 44, no. 11–12, pp. 732–742, 2006.
- [19] S. Weeda *et al.*, "Arabidopsis transcriptome analysis reveals key roles of melatonin in plant defense systems," *PLoS One*, vol. 9, no. 3, p. e93462, 2014.
- [20] U. Nagalakshmi *et al.*, "The transcriptional landscape of the yeast genome defined by RNA sequencing," *Science (80-)*, vol. 320, no. 5881, pp. 1344–1349, 2008.
- [21] H. Li *et al.*, "The sequence alignment/map format and SAMtools," *Bioinformatics*, vol. 25, no. 16, pp. 2078–2079, 2009.
- [22] A. hub web presence working group, "AmiGO: online access to ontology and annotation data," *Bioinformatics*, vol. 25, 2009.
- [23] M. Kanehisa and S. Goto, "KEGG: kyoto encyclopedia of genes and genomes," *Nucleic Acids Res.*, vol. 28, no. 1, pp. 27–30, 2000.
- [24] T. Altman, M. Travers, A. Kothari, R. Caspi, and P. D. Karp, "A systematic comparison of the MetaCyc and KEGG pathway databases," *BMC Bioinformatics*, vol. 14, no. 1, pp. 1–15, 2013.
- [25] P. D. Karp, M. Riley, S. M. Paley, and A. Pellegrini-Toole, "The metacyc database," *Nucleic Acids Res.*, vol. 30, no. 1, pp. 59–61, 2002.
- [26] Z. Pang *et al.*, "MetaboAnalyst 5.0: narrowing the gap between raw spectra and functional insights," *Nucleic Acids Res.*, vol. 49, no. W1, pp. W388–W396, 2021.
- [27] S. Naithani *et al.*, "Plant Reactome: a knowledgebase and resource for comparative pathway analysis," *Nucleic Acids Res.*, vol. 48, no. D1, pp. D1093–D1103, 2020.
- [28] J. SONG *et al.*, "Mining genes associated with furanocoumarin biosynthesis in an endangered medicinal plant, *Glehnia littoralis*," *J. Genet.*, vol. 99, no. 1, p. 11, 2020, doi: 10.1007/s12041-019-1170-6.
- [29] Y. Yang *et al.*, "Transcriptome analysis of *Aconitum carmichaelii* and exploration of the salsolinol biosynthetic pathway," *Fitoterapia*, vol. 140, p. 104412, 2020.
- [30] S. Zhang, L. Zhao, C. Shan, Y. Shi, K. Ma, and J. Wu, "Exploring the biosynthetic pathway of lignin in *Acorus tatarinowii* Schott using de novo leaf and rhizome transcriptome analysis," *Biosci. Rep.*, vol. 41, no. 6, 2021.
- [31] Z. Zhang *et al.*, "Integrative transcriptomics and metabolomics data exploring the effect of chitosan on postharvest grape resistance to *Botrytis cinerea*," *Postharvest Biol. Technol.*, vol. 167, p. 111248, 2020.
- [32] N. R. Council, "Committee on Toxicity Testing and Assessment of Environmental Agents. Toxicity testing in the 21st century: a vision and a strategy." Washington, DC: National Academies Press, 2007.
- [33] D. Krewski *et al.*, "Toxicity testing in the 21st century: a vision and a strategy," *J. Toxicol. Environ. Heal. Part B*, vol. 13, no. 2–4, pp. 51–138, 2010.
- [34] J. K. Padhan, P. Kumar, H. Sood, and R. S. Chauhan, "Prospecting NGS-transcriptomes to assess regulation of miRNA-mediated secondary metabolites biosynthesis in *Swertia chirayita*, a medicinal herb of the North-Western Himalayas," *Med. Plants-International J. Phytomedicines Relat. Ind.*, no. 3, pp. 219–228, 2016.
- [35] N. I. Maulidah, T.-S. Tseng, G.-H. Chen, H.-Y. Hsieh, S.-F. Chang, and H. Chuang, "Transcriptome analysis revealed cellular pathways associated with abiotic stress tolerance and disease resistance induced by *Pseudomonas aeruginosa* in banana plants," *Plant Gene*, vol. 27, p. 100321, 2021.
- [36] F. Cartieaux *et al.*, "Transcriptome analysis of *Arabidopsis* colonized by a plant-growth promoting rhizobacterium reveals a general effect on disease resistance," *Plant J.*, vol. 36, no. 2, pp. 177–188, 2003.