Transcriptome Analysis: Connecting the Genome to Gene Function

Sainath K[1]*, Praveen K[3], B P Maruti Prasad[4], Eshwar[1] and Akshaya M[2]

[1]Department of Genetics and Plant Breeding, UAS, Raichur
[2]Department of Genetics and Plant Breeding, UAS, Dharwad
[3]Department of Agricultural Microbiology, UAS, Bangalore
[4]Department of Genetics and Plant Breeding, UAS, Bangalore

Corresponding Author
Sainath K
Email: sainathkarbhari98@gmail.com

Keywords

Non-coding RNAs, non-coding RNAs, transcriptome analysis, Next-generation sequencing technology and Transcription profiling

How to cite this article:

Sainath, K., Praveen, K., Prasad, B. P. M., Eshwar. and Akshaya, M. 2025. Transcriptome Analysis: Connecting the Genome to Gene Function. *Vigyan Varta* 6 (12): 39-43.

ABSTRACT

Transcriptomics, the study of the complete set of RNA transcripts produced by the genome, has become a central tool in the post-genomic era for understanding gene expression and regulation. It provides critical insights into cellular functions, developmental processes and responses to environmental or physiological stimuli. Techniques such as microarrays and RNA-sequencing (RNA-Seq) have revolutionized transcriptome profiling by enabling comprehensive quantification of both coding and non-coding RNAs. These methods not only identify novel transcripts and splicing variants but also uncover expression patterns underlying disease mechanisms, stress tolerance and developmental regulation. Transcriptome data further support integrative analyses in systems biology, linking gene activity to protein networks and metabolic pathways. Thus, transcriptome analysis serves as a foundational approach for discovering biomarkers, functional genes and therapeutic targets, significantly advancing personalized medicine and crop improvement.

INTRODUCTION

With the advent of the post-genomic era, various omics techniques such as transcriptomics, proteomics, and metabolomics have emerged, of which transcriptomics is the first and most widely used technology. Transcriptomics studies transcriptome. A transcriptome is the sum of all RNAs transcribed from particular tissue or cell at a certain developmental stage or functional state, including messenger RNA (mRNA) and non-coding RNA (ncRNA). mRNA, considered as 'bridges', precisely regulates the transmission of genetic information from DNA to protein, while ncRNA regulates gene expression, protein synthesis, and different cellular activities at different levels. Therefore, understanding of transcriptome promotes the study of the activities of cells, tissues and organisms (Zhang, 2019).

Transcriptome sequencing, also known as RNA-Seq, is a recently developed technique that quantifies transcriptome in given samples and thus promotes transcriptome analysis. The key aims of transcriptomics are: to catalogue all species of transcript, including mRNAs, non-coding RNAs and small RNAs; to determine the transcriptional structure of genes, in terms of their start sites, splicing patterns and other post-transcriptional modifications and to quantify the changing expression levels of each transcript during development and under different conditions. After the genome has been sequenced, transcriptome analysis allows us to understand the expression of genome at the transcription level, which provides information on gene structure, regulation of gene expression, gene product function and genome dynamics.

Next-generation sequencing technology is a powerful and cost-efficient tool for ultra-high-throughput transcriptome analysis. By analyzing the transcriptome at spectacular and unprecedented depth and accuracy, thousands of new transcript variants and isoforms have been shown to be expressed in mammalian tissues or organs. These advances greatly accelerate our understanding of the complexity of gene expression, regulation and networks for mammalian cells. These new techniques usually need microgram amounts of total RNA for analysis, which corresponds to hundreds of thousands of mammalian cell (Patwekar *et al.,* 2025).

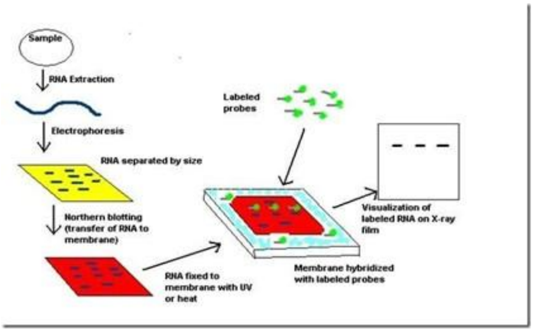
A major challenge in the molecular biology continues to be the complex mapping of the same genome to diverse phenotypes in different tissue types, development stages and environmental conditions. A better understanding of the transcripts and expression of gene regulation is not only non-trivial but lies at the heart of this challenge. Transcriptomics offers important insights on gene structure, expression and regulation and has been widely studied in many organisms. The transcriptomics studies have advanced considerably because of the explosive growth in the underlying sequencing technology.

Transcription profiling, which follows changes in behavior of a cell in to, not of a single gene or just a few genes, is used throughout diverse areas of biomedical research, including disease diagnosis, biomarker discovery, risk assessment of new drugs or environmental chemicals *etc.* Transcription profiling can be applied to loss- and gain-of-function mutants to identify the changes associated with the mutant phenotype. The transcriptomic techniques have been particularly useful in identifying the functions of genes. Transcriptomics also allows identification of pathways that respond to or ameliorate environmental stresses. RNA-Seq can also identify disease-associated gene fusions, single nucleotide polymorphisms and even allele-specific expression (Blumenberg, 2019).

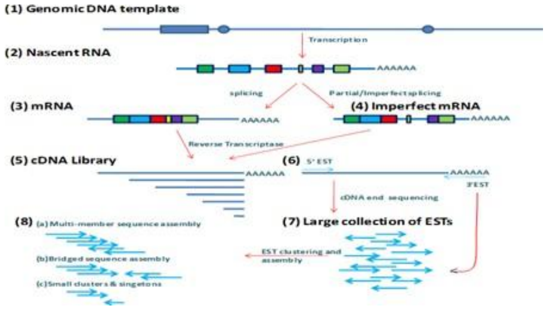
TRANSCRIPTOME SEQUENCING TECHNIQUES

There are so many sequencing techniques were developed to analyse the gene expression at transcriptome level or mRNA level to understand the mechanisms involved in final phenotypic expression. Those techniques were listed below (Lowe *et al.,* 2017).

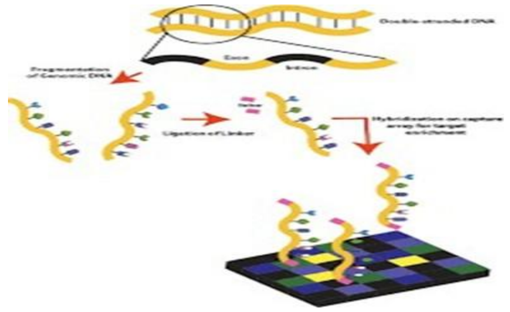
1. Northern blotting



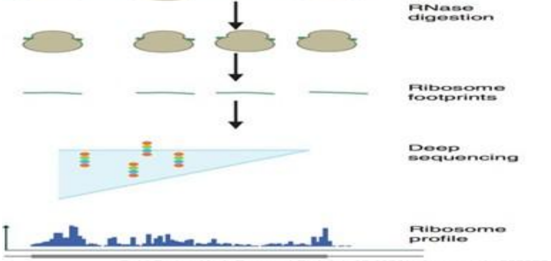
2. Expressed sequence tags (ESTs)



3. Exome Sequencing

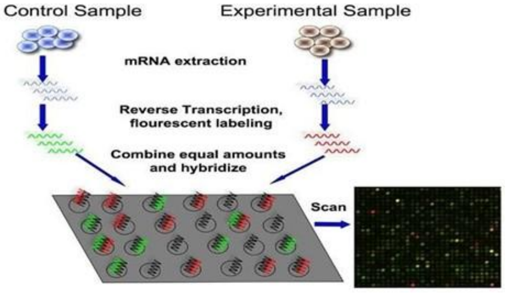


4. Ribosome profiling

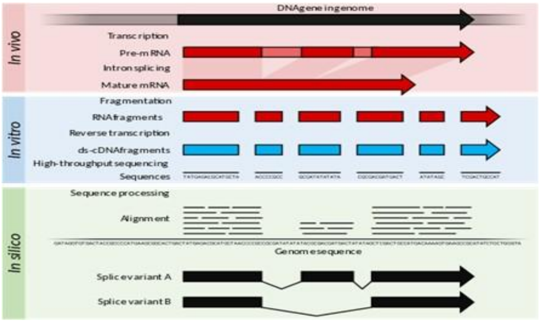


Cold Spring Harb Perspect Biol doi: 10.1101/cshperspect.a032698

5 Microarrays



6 RNA-Sequencing



Steps involved in RNA-Sequencing:

1. Sample preparation
2. transcript enrichment
3. Library preparation
   i. Fragmentation
   ii. Amplification
   iii. Single or paired-end sequencing (Kukurba & Montgomery, 2015)

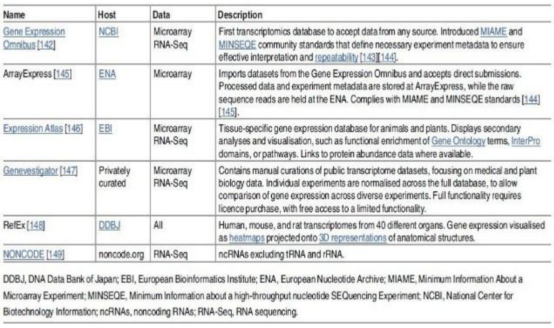
Steps involved in transcriptome data analysis are as follows

- Raw reads generation from the sequence
- Raw sequence reads quality control
- Mapping of raw reads
- Counting
- Differential analysis
- Biological interpretation

Transcriptome Databases:

- Transcriptomics studies generate large amounts of data that has potential applications far beyond the original aims of an experiment.

As such, raw or processed data may be deposited into public databases to ensure their utility for the broader scientific community.



Name	Host	Data	Description
Gene Expression Omnibus [142]	NCBI	Microarray RNA-Seq	First transcriptomics database to accept data from any source. Introduced MIAME and MINSEQE community standards that define necessary experiment metadata to ensure effective interpretation and repeatability [143][144].
ArrayExpress [145]	ENA	Microarray	Imports datasets from the Gene Expression Omnibus and accepts direct submissions. Processed data and experiment metadata are stored at ArrayExpress, while the raw sequence reads are held at the ENA. Complies with MIAME and MINSEQE standards [144][145].
Expression Atlas [146]	EBI	Microarray RNA-Seq	Tissue-specific gene expression database for animals and plants. Displays secondary analyses and visualisation, such as functional enrichment of Gene Ontology terms, InterPro domains, or pathways. Links to protein abundance data where available.
Genevestigator [147]	Privately curated	Microarray RNA-Seq	Contains manual curations of public transcriptome datasets, focusing on medical and plant biology data. Individual experiments are normalised across the full database, to allow comparison of gene expression across diverse experiments. Full functionality requires licence purchase, with free access to a limited functionality.
RefEx [148]	DDBJ	All	Human, mouse, and rat transcriptomes from 40 different organs. Gene expression visualised as heatmaps projected onto 3D representations of anatomical structures.
NONCODE [149]	noncode.org	RNA-Seq	ncRNAs excluding tRNA and rRNA.

DDBJ, DNA Data Bank of Japan; EBI, European Bioinformatics Institute; ENA, European Nucleotide Archive; MIAME, Minimum Information About a Microarray Experiment; MINSEQE, Minimum Information about a high-throughput nucleotide SEQuencing Experiment; NCBI, National Center for Biotechnology Information; ncRNAs, noncoding RNAs; RNA-Seq, RNA sequencing.

Role and Applications of Transcriptome Analysis in Modern Biology (Stahl et al., 2011)

☐ **Comprehensive Gene Activity Profiling:**

- Enables characterization of the complete set of RNA transcripts (coding and non-coding) expressed in a cell or tissue.
- Helps create a global picture of cellular function and regulatory mechanisms.

☐ **Gene Expression Analysis:**

- Provides a snapshot of actively expressed genes and transcripts under specific conditions.
- Commonly used to compare samples under different environmental, developmental, or physiological conditions (e.g., hormone treatment, toxin exposure, or stress).

☐ **Disease vs. Healthy Comparison:**

- Widely applied to compare healthy and diseased tissues to identify differentially expressed genes.
- Useful in understanding disease mechanisms, pathogenesis, and identifying diagnostic biomarkers.

☐ **Applications in Cancer Research:**

- Helps classify cancer types beyond anatomical and histopathological criteria.
- Assists in understanding tumor biology, mechanisms of carcinogenesis, and identifying therapeutic targets.
- Facilitates **outcome prediction** and **personalized medicine**, including prognosis and therapy response based on gene expression profiles.

☐ **Developmental and Stem Cell Studies:**

- Used to characterize gene expression patterns during different developmental stages.
- Transcriptome analysis of stem cells helps understand **cellular differentiation** and **embryonic development** processes.

☐ **Environmental and Stress Response Studies:**

- Assesses transcriptional changes under various environmental stresses (e.g., drought, salinity, temperature).
- Identifies genes involved in adaptive and stress response pathways.

☐ **Identification of Novel Targets:**

- Provides a broad dataset for discovering **new genes**, **biomarkers**, and **molecular targets** for disease treatment and crop improvement.
- Supports the development of gene-based diagnostic tools and therapeutic strategies.

☐ **Systems Biology Integration:**

- Acts as a foundation for systems-level understanding by linking gene expression to protein function and metabolic pathways.
- Enhances integrative research combining genomics, proteomics, and metabolomics.

CONCLUSION

Transcriptome analysis provides a powerful platform to explore the functional complexity of genomes by revealing when, where, and how genes are expressed. With the rapid advancement of next-generation sequencing technologies, transcriptomics enables high-resolution profiling of gene expression across tissues, developmental stages and environmental conditions. It bridges the gap between genomic information and phenotypic traits, facilitating discoveries in disease diagnostics, stress biology and functional genomics. As transcriptome databases expand and analytical tools improve, this field continues to enhance our understanding of molecular mechanisms and gene regulation. Ultimately, transcriptome analysis has become indispensable in modern biology for translating genomic data into biological meaning and practical applications in medicine and agriculture.

REFERENCES

Blumenberg, M. (2019). *Introductory chapter: Transcriptome analysis*. IntechOpen.

Kukurba, K. R., & Montgomery, S. B. (2015). RNA sequencing and analysis. *Cold Spring Harbor Protocols*, 2015(11): 084970.

Lowe, R., Shirley, N., Bleackley, M., Dolan, S., & Shafee, T. (2017). Transcriptomics technologies. *PLoS Computational Biology*, 13(5): e1005457.

Patwekar, M., Patwekar, F., Badarinath, A. V., Billah, A. A. M., Gorijavolu, V., Krishnan, K., Shanmugasundaram, P., Prasad, P. D., & Kazi, A. A. (2025). Genomic sequencing: Techniques, advancements, and the path ahead. *Journal of Bio-X Research*, 8: 0046.

Stahl, F., Hitzmann, B., Mutz, K., Landgrebe, D., Lübbecke, M., Kasper, C., Walter, J., & Scheper, T. (2011). Transcriptome analysis. In *Genomics and Systems Biology of Mammalian Cell Culture* 1: 1–25.

Zhang, H. 2019. The review of transcriptome sequencing: Principles, history and advances. In *IOP Conference Series: Earth and Environmental Science*, 332 (4): 042003.