The Pharma Innovation Journal 2023; SP-12(11): 391-394

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; SP-12(11): 391-394

www.thepharmajournal.com
Received: 01-08-2023
Accepted: 05-09-2023

Raju Ram
ICAR - Central Institute of Fisheries Education, Panch Marg, Yari Road, Mumbai, Maharashtra, India

Gowhar Iqbal
ICAR - Central Institute of Fisheries Education, Panch Marg, Yari Road, Mumbai, Maharashtra, India

Jaspreet Singh
Scientist, Division of livestock and Fisheries Management, ICAR Research Complex for Eastern Region, Patna, Bihar, India

Devarshi Ranjan
ICAR-National Bureau of Fish Genetic Resources, Canal Ring Road, P.O. Dilkusha, Lucknow, Uttar Pradesh, India

Vikas Pathak
Assistant Professor, College of Fisheries, The Neotia University, 24 Parganas (South), West Bengal, India

Shyam Kumar
Scientist, KVK, Sheohar, Bihar, India

Amrita Mohanty
ICAR - Central Institute of Fisheries Education, Panch Marg, Yari Road, Mumbai, Maharashtra, India

Corresponding Author:
Raju Ram
ICAR - Central Institute of Fisheries Education, Panch Marg, Yari Road, Mumbai, Maharashtra, India

RNA sequencing: A technique for advancing the field of fish genetics and breeding

Raju Ram, Gowhar Iqbal, Jaspreet Singh, Devarshi Ranjan, Vikas Pathak, Shyam Kumar and Amrita Mohanty

Abstract
Global gene expression profiling using RNA-seq, enabled by high throughput third-generation sequencing technology, revolutionizes transcriptome exploration. Compared to microarray, RNA-seq has advantages, including a larger dynamic range, accurate transcript quantification, and comprehensive tissue expression coverage. In addition to quantification of the transcript, RNA-seq is used for transcriptome comparison, mapping the both 5' and 3' ends of introns and mRNA, discovery of new exons, research into allele-specific expression, SNP finding, construction of de novo gene models and investigation into non-coding RNA. RNA-seq technology is increasingly utilized in various commercial and ecologically important species-related research for differential gene expression study and SNP discovery. The present review article encapsulates the RNA-seq profiling steps and potential applications in aquaculture.

Keywords: RNA-seq, differential gene expression, SNP, aquaculture, transcriptome

Introduction
Transcriptomics investigates the entire collection of transcripts in a particular cell, tissue, or organism for a specific developing stage or functional state [1]. A transcriptome is a comprehensive collection of transcripts that includes both messenger RNA (mRNA) and non-coding RNA (ncRNA), such as transfer RNA (tRNA), ribosomal RNA (rRNA), and other ncRNAs). Unlike the stable genome, transcriptome changes based on the development stage, body state, and environment. Transcriptome analysis helps understand cell fate, growth, and illness development by examining the relationship between genotype and phenotype. It maps and annotates the transcriptome, identifies gene functional organization, and quantifies gene expression changes across different samples [1, 2]. High-throughput technology is needed for transcriptome research due to its complexity, and methods like hybridization and sequence-based techniques have been developed. Hybridization-based technologies involve incubating fluorescently labelled cDNA with probes on microarrays. RNA-seq is an affordable technique for mapping and quantifying transcriptomes that uses NGS technology. RNA-seq and other technologies analyze comprehensive personal omics profiles for disease progression monitoring and risk assessment to provide individualized therapy [3]. This review discusses the RNA-seq method's overview, challenges, and biological insights gained for various fish species, focusing on data generation and bioinformatic analysis.

Fundamentals of RNA-seq
RNA-seq sequences cDNA directly from a sample using next-generation sequencing technology [4]. For transcriptome analysis, RNA-seq requires library creation, sequencing, and bioinformatics (Figure 1).

1. Library preparation
Library construction is crucial for accurate RNA-seq data resemblance to the transcriptome. The process begins with collecting samples, often tissues, for analysis. Then, total RNA is produced using silica membrane absorption or organic extraction. Whole RNA samples undergo processing through ribo- depletion or direct selection of poly (A) RNA, using an oligo(dT)-based purification method for eukaryotic mRNA [2]. Then, a double-stranded cDNA library is created through reverse transcription followed by DNase I treatment to fragment the cDNA or RNA fragmentation [1, 5].

Further, these cDNA fragments are ligated with sequencing adapters at both ends.

2. High throughput Sequencing

NGS technologies predominate in the sequencing industry, and RNA-seq extensively uses popular platforms, including Roche's 454 pyrosequencing system, Life Technologies' AB SOLiD system, and Illumina's Illumina Genome Analyzer [6]. These NGS systems use *in vitro* cloning for amplifying fragmented cDNA in a cell-free environment, as their sensitivities are insufficient for single-molecule sequencing [7]. Illumina produces longer paired-end reads that are more effective for de novo assemblies. In brief, Illumina uses oligonucleotide hybridization to immobilize single-stranded cDNA fragments on a glass slide to carry out clonal PCR amplification. The advanced generation sequencing methods, including Heliscope sequencing and Single-Molecule Real-Time (SMRT), offer real-time signal capture and template clonal amplification [6].

3. Bioinformatic analysis

NGS sequencers generate millions of raw reads, providing base-call quality ratings to assess the reliability of each call. Quality assessment is crucial for RNA-seq data analysis, removing adapter sequences, poor-quality reads, and overrepresentation to ensure a coherent final result. FastQC and Trimmomatic are such software that can assist in this process. High-quality small sequencing reads are aligned to a standard reference genome using popular programs like ELAND, SOAP, MAQ, and RMAP [1]. In the absence of the reference genome or transcriptome, de novo transcriptome assembly is carried out. De novo assembly software like ABySS, Velvet, and Trinity utilize the de Bruijn graph approach, aligning sequence overlap (k-mer) between reads to generate contigs [8]. After mapping reads, transcript expression levels can be estimated by counting exons in a gene. This generates a genome-scale transcriptome map incorporating quality and quantity by assigning each base an expression score. RNA-seq analysis also identifies molecular markers, fusion genes, and post-transcriptional gene control, including RNA editing, depletion, and translations [9].

Potential usage of RNA-seq in aquaculture

RNA-seq in aquaculture aids in identifying novel transcripts, molecular marker detection, quantifying gene expressions and generating genetic resources.

1. Novel transcript findings

RNA-seq technology reveals novel transcribed areas in fish species like zebrafish [10], rohu carp [11], and rainbow trout [12], indicating a significant number of unidentified transcribed areas. Among these unidentified transcripts, several non-coding RNAs (ncRNAs) other than rRNA and tRNA are of major interest due to their critical participation in some cellular processes, including RNA splicing, protein translation, regulation, and genome protection. [13]. These newly discovered transcripts aid genome annotation and functional studies in future research. Deep RNA sequencing was used to identify genes and quantify the effects of exercise on the transcriptomes of skeletal muscles in both exercised and non-exercised rainbow trout. The research discovered 1,085 and 1,228 new rainbow trout gene sequences, which indicates that continuous swimming enhances transcriptional movement in skeletal muscle tissues and up regulates transcripts related to muscular development and growth [14].

2. Molecular marker identification

SNPs are single nucleotide variations (SNPs) found in the genome, used as markers in various genomic and genetic applications. RNA-seq is a successful method for detecting gene-associated SNPs genome-wide due to its sensitivity and single-base precision. These markers are often employed in genome-wide association studies and quantitative QTL mapping [15]. There have already been several reports of SNP identification-directed RNA-seq investigations in fish for significant traits such as growth [16], immunity [17], reproduction [11], and stress [18].

3. Transcript level quantification

RNA-seq is a quantitative technique that records transcriptome dynamics in response to external or internal factors, providing accurate assessments of transcript levels. It is widely used in fish research for developmental biology, immunology, evolutionary biology, physiology, toxicity, and illnesses to identify differential gene expression. For instance, the research investigated the differential expression of immune-associated genes in gibel carp (*Carassius auratus gibelio*) following infection with pharyngeal myxobolosis [19]. A study analyzed grass carp muscle tissue transcriptomes, revealing 35 down-regulated and 71 up-regulated genes in DEGs involved in muscle tissue growth. Up-regulation of IGFBP1, GHR2, ALAS2, and myoglobin 1 may encourage dietary intake, aerobic activity, and muscular development [20]. These studies reveal key genes for fish biomarkers, enhancing understanding molecular mechanisms and regulatory pathways.

4. Genetic resource creation

RNA-seq facilitates the discovery of genes related to economic aspects, genes involved in overall growth, and reproduction aspects such as gonad differentiation, sex determination, and immune response. RNA sequences improve the quality of genetic resources and hence increase the size of the database. Additionally, it makes genetic approaches for controlling sex ratios, enhancing growth, and strengthening disease resistance possible. De novo transcriptome assembly from fish species offers genetic data for studies on development, immunity, and other physiological processes. Lumpfish or lumpsucker (*Cyclopterus lumpus*) transcriptome profiling reveals transcripts related to immune response, growth, and molecular markers [21]. Challenge trials could scan the transcriptome to examine alterations in mechanism and immunological reactions to *V. anguillarum* in Half-smooth tongue sole (*Cynoglossus semilaevis*) [22].

5. Molecular pathway assessment

Gene ontology studies define cellular components, pathways, and processes in RNA-seq analysis, describing molecular functions and events of various genetic produces. For instance, the study examined salinity adaptation in Nile tilapia, examining its steroid, lipid, immunological, and osmoregulation pathways [23]. The Mozambique tilapia's gill transcriptome was studied to understand its euryhaline nature after acclimatization [24]. This study examined signalling pathways related to cell cycle, division, DNA replication, metabolism, and protein synthesis to control hyperosmotic conditions. Hu *et al*. [25] performed liver transcriptome

analysis on *Oncorhynchus mykiss* to find new genes related to biological indicators of fat accumulation and lipid absorption.

6. Single-cell RNA sequencing

Analyzing each transcriptome helps understand functional heterogeneity within a population due to gene expression variations and network inference. This situation has recently undergone a significant evolution made possible by single-cell mRNA sequencing (scRNA-seq) techniques. These methods enable researchers to study large numbers of genes, revealing the influence of biological cues on cell responses and the impact of epigenetic changes like DNA methylation and histone modification on each group of genes [15]. Transcriptome analysis using RNA sequencing technology offers high-resolution insights into cell state, destiny, and variability in genetically homogeneous individuals [26]. Researchers conducted the single-cell transcriptomics-based characterization of immune cell repertoire in syngnathid fish (*Syngnathus typhle*) [27]. This study lays the groundwork for future cellular categorization and experimental work within the lineage. A transcriptomic atlas from Atlantic salmon reveals liver cellular heterogeneity in fish, revealing poor conservation of hepatic cell marker genes and novel heterogeneity within hepatocyte, lymphoid, and myeloid lineages [28].

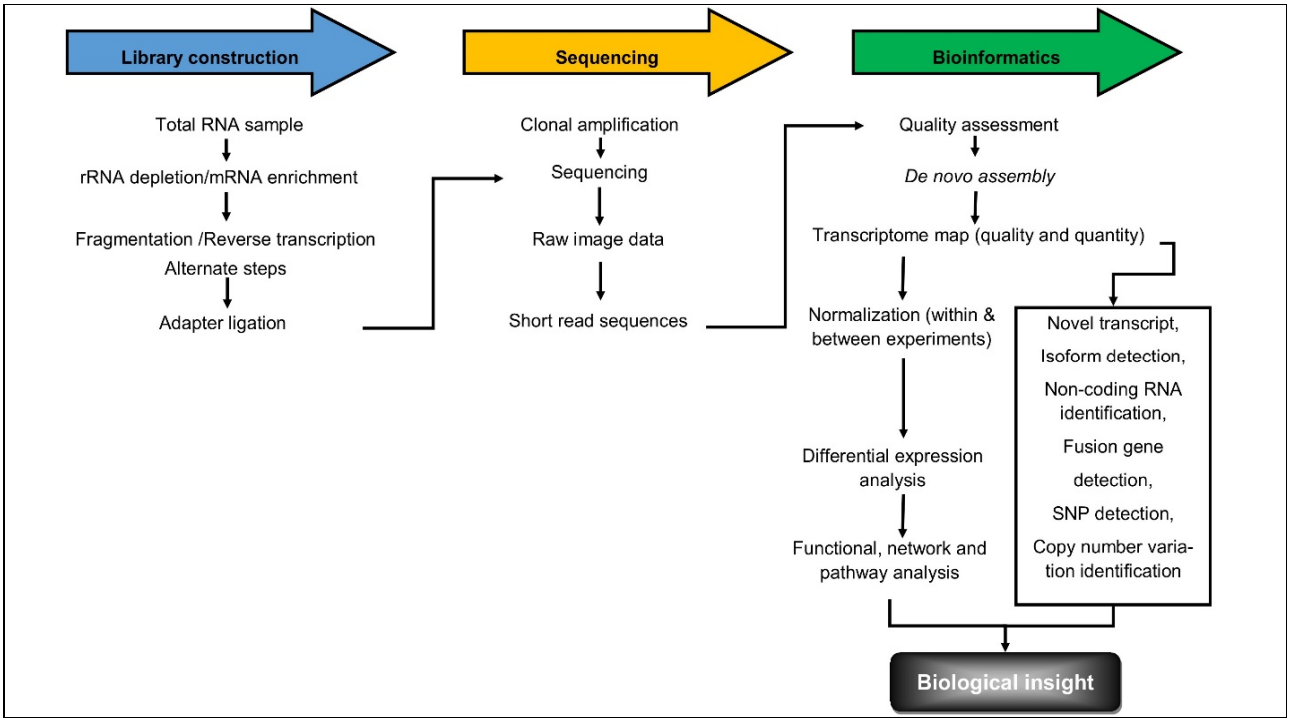

Fig 1: Illustration of the workflow for analyzing RNA-Seq data

Conclusion

The advent of the RNA-seq approach in transcriptomics has led to an explosion of de novo transcriptome data, making it essential for many genomic studies. RNA-seq is crucial for long-term sustainability in aquaculture, as it makes it possible to identify genes expressed differently in test organisms than controls. This technology can also create molecular markers linked to significant aquaculture features. ScRNA-seq technology enhances overall immune responses in fishes by revealing immune cell contributions and interactions. This analysis provides an overall comprehensive analysis of RNA Seq's current state and use in aquaculture.

References

1. Wang Z, Gerstein M, Snyder M. RNA-Seq: A revolutionary tool for transcriptomics. Nat Rev Genet. 2009;10(1):57-63.
2. Costa V, Angelini C, Feis DI, Ciccodicola A. Uncovering the complexity of transcriptomes with RNA-Seq. Biomed Res Int; c2010.
3. Chen R, Mias GI, Li-Pook-Than J, Jiang L, Lam HY, Chen R, *et al*. Personal omics profiling reveals dynamic molecular and medical phenotypes. Cell. 2012;148(6):1293-1307.
4. Morozova O, Hirst M, Marra MA. Applications of new sequencing technologies for transcriptome analysis. Annu Rev Genom Hum Genet. 2009;10:135-51.
5. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol. 2011;12(3):1-4.
6. Liu L, Li Y, Li S, Hu N, He Y, Pong R, *et al*. Comparison of next-generation sequencing systems. J Biomed Biotechnol; c2012.
7. Metzker ML. Sequencing technologies-the next generation. Nat Rev Genet. 2010;11(1):31-46.
8. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, *et al*. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29(7):644-652.
9. Marguerat S, Bähler J. RNA-seq: from technology to biology. Cell Mol Life Sci. 2010;67:569-579.
10. Wang J, Vesterlund L, Kere J, Jiao H. Identification of novel transcribed regions in zebrafish (*Danio rerio*) using RNA-sequencing. PLoS One. 2016;11(7):e0160197.
11. Jaiswal S, Nandi S, Iquebal MA, Jasrotia RS, Patra S, Mishra G, *et al*. Revelation of candidate genes and

molecular mechanism of reproductive seasonality in female rohu (*Labeo rohita* Ham.) by RNA sequencing. BMC Genom. 2021;22(1):1-21.
12. Sun J, Liu Z, Quan J, Li L, Zhao G, Lu J. RNA-seq analysis reveals alternative splicing under heat stress in rainbow trout (*Oncorhynchus mykiss*). Mar Biotechnol. 2022;1:1-3.
13. Mattick JS, Makunin IV. Non-coding RNA. Hum Mol Genet. 2006;15(suppl_1):R17-29.
14. Palstra AP, Beltran S, Burgerhout E, Brittijn SA, Magnoni LJ, Henkel CV, *et al*. Deep RNA sequencing of the skeletal muscle transcriptome in swimming fish. PLoS One. 2013;8(1):e53171.
15. Chandhini S, Kumar RVJ. Transcriptomics in aquaculture: current status and applications. Rev Aquac. 2019;11(4):1379-1397.
16. Guan WZ, Qiu GF. Transcriptome analysis of the growth performance of hybrid mandarin fish after food conversion. PLoS One. 2020;15(10):e0240308.
17. Gervais O, Barria A, Papadopoulou A, Gratacap RL, Hillestad B, Tinch AE, *et al*. Exploring genetic resistance to infectious salmon anaemia virus in Atlantic salmon by genome-wide association and RNA sequencing. BMC Genom. 2021;22(1):1-4.
18. Liu Y, Jiang T, Chen Y, Gu Y, Song F, Sun J, *et al*. Identification of Candidate Genes Associated With Hypoxia Tolerance in *Trachinotus blochii* Using Bulked Segregant Analysis and RNA-Seq. Front Genet. 2021;12:811685.
19. Zhao Y, Liu X, Sato H, Zhang Q, Li A, Zhang J. RNA-seq analysis of local tissue of *Carassius auratus gibelio* with pharyngeal myxobolosis: Insights into the pharyngeal mucosal immune response in a fish-parasite dialogue. Fish Shellfish Immunol. 2019;94:99-112.
20. Lu X, Chen HM, Qian XQ, Gui JF. Transcriptome analysis of grass carp (*Ctenopharyngodon idella*) between fast-and slow-growing fish. Comp Biochem Physiol-D: Genom Proteom. 2020;35:100688.
21. Maduna SN, Vivian-Smith A, Jónsdóttir ÓD, Imsland AK, Klütsch CF, Nyman T, *et al*. Genome-and transcriptome-derived microsatellite loci in lumpfish *Cyclopterus lumpus*: molecular tools for aquaculture, conservation and fisheries management. Sci Rep. 2020;10(1):559.
22. Qi L, Chen Y, Shi K, Ma H, Wei S, Sha Z. Combining of transcriptomic and proteomic data to mine immune-related genes and proteins in the liver of *Cynoglossus semilaevis* challenged with *Vibrio anguillarum*. Comp Biochem Physiol-D: Genom Proteom. 2021;39:100864.
23. Xu Z, Gan L, Li T, Xu C, Chen K, Wang X, *et al*. Transcriptome profiling and molecular pathway analysis of genes in association with salinity adaptation in Nile tilapia *Oreochromis niloticus*. PLoS One. 2015;10(8):e0136506.
24. Lam SH, Lui EY, Li Z, Cai S, Sung WK, Mathavan S, *et al*. Differential transcriptomic analyses revealed genes and signaling pathways involved in iono-osmoregulation and cellular remodeling in the gills of euryhaline Mozambique tilapia, *Oreochromis mossambicus*. BMC Genom. 2014;15(1):1-23.
25. Hu P, Liu M, Liu Y, Wang J, Zhang D, Niu H, *et al*. Transcriptome comparison reveals a genetic network regulating the lower temperature limit in fish. Sci Rep. 2016;6(1):28952.
26. Liu S, Trapnell C. Single-cell transcriptome sequencing: recent advances and remaining challenges. F1000 Research, 2016, 5.
27. Parker J, Guslund NC, Jentoft S, Roth O. Characterization of pipefish immune cell populations through single-cell transcriptomics. Front Immunol. 2022;13:131.
28. Taylor RS, Daniels RR, Dobie R, Naseer S, Clark TC, Henderson NC, *et al*. Single-cell transcriptomics of Atlantic salmon (*Salmo salar* L.) liver reveals cellular heterogeneity and immunological responses to challenge by *Aeromonas salmonicida*. Front Immunol. 2022;13:984799.