International Journal of Agriculture Sciences
ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 14, Issue 12, 2022, pp.-12006-12009.
Available online at https://bioinfopublication.org/pages/jouarchive.php?id=BPJ0000217

Review Article

RNA-SEQUENCE AS AN AGRICULTURAL SCIENCE RESEARCH BOOSTER

BANIK M.[1]*, MANDAL A.K.[2], NILANJAYA[1], SHAH P.[3] AND KUMAR J.[4]

[1]Department of Plant Breeding and Genetics, Dr Rajendra Prasad Central Agricultural University, Pusa, 848125, Bihar, India
[2]Department of Agrometeorology, Dr Kalam Agricultural College, Arrabari, Kishanganj, 855107, Bihar Agricultural University, Sabour, 813210, Bihar, India
[3]Department of Genetics and Plant Breeding, SASRD, Nagaland University, Medziphema, 797106, Nagaland, India
[4]Department of Agriculture Engineering, Dr Kalam Agricultural College, Arrabari, Kishanganj, 855107, Bihar Agricultural University, Sabour, 813210, Bihar, India
*Corresponding Author: Email - mahi.jit92@gmail.com

Received: December 05, 2022; Revised: December 26, 2022; Accepted: December 28, 2022; Published: December 30, 2022

Abstract: RNA- sequence is a revolutionized next generation sequencing technology. This analysis does not demand gene probes rather provide an accurate measure of gene expression as compared to other commonly available techniques. The transcription and gene profile data can be store in database and could be accessed by researcher. Transcriptome is the collective information of all the gene readouts present in a cell including coding and non-coding RNA transcript. Genes that are differentially expressed during different developmental stages can be identified by aspect of transcriptomics. The fact that RNA sequencing does not entails gene probes rather provides precise measures of gene expression over a much wider range, which has demonstrated its credibility over other common technique. Thus, RNA sequencing gives very crucial genetic information along with alternative splicing and isoforms of genes.

Keywords: *RNA sequence method, High throughput sequencing, Expressed gene analysis*

Citation: Banik M., *et al*., (2022) RNA-Sequence as an Agricultural Science Research Booster. International Journal of Agriculture Sciences, ISSN: 0975-3710 & E-ISSN: 0975-9107, Volume 14, Issue 12, pp.- 12006-12009.

Academic Editor / Reviewer: Altino Choupina

Introduction

Agricultural science, plant breeding is the branch which made considerable effort to breed new varieties. But with the burgeoning population there is huge demand of higher food grain production so how does an agriculturist will fulfil this challenge? Therefore, an immediate need is felt to co-integrate conventional breeding tool with biotechnology to speed up the crop improvement programs. Several biotechnological methods can be useful to have better ones in the process of plant breeding [1]. The magic method of RNA sequencing helps to understand which genes involved in expression of various mechanism of cells and organs of a plant body. Recently in this arena many research articles have been published.

RNA-sequencing (RNA-seq) technology utilizes Next generation sequencing (NGS) principle and it was first used to study plants only a few years ago [2]. It provides ready access to high resolution transcriptome information to an extent that was once unimaginable [3]. It is the high throughput sequencing of complementary DNA (cDNA) fragment populations. RNA- sequencing is the most rapidly used technology in recent year; due to accuracy in analysis and determination at transcription level as well as for reproducibility and precision [4]. Thus, it has a wide spectrum application in the plant science research. It has been already successfully applied in transcriptome profiling of species without plants of known genomes sequence [5, 6].

In principle in RNA-Sequencing process first the mRNA need to be reverse transcribe into cDNA, after that attach the splice (adapter) and then place to a sequencer to be sequenced. The resulting sequence reads are aligned to a reference transcript or, assembled as de novo to produce a genome scaled map consisting of transcript structure and its expression level for each gene specific stage of development. RNA-Sequencing widely applied to model along with non-model individual to get mass sequence data for molecular marker development, analysis of transcription for discovery of novel gene, especially in examination transcriptome fine structure such as defution of alleles specific expiration and variation at splice junction [7].

In recent trends, a four high throughput sequencing technology Rocho454, Illumine, Life technologies and Helicos bio sciences are used. And still other new modern technologies are in queue of development process. Although illumine system is rapidity used for it cheaper cost per base sequence.

Instances of RNA-Sequencing in plant breeding system

In regards to plant breeding system RNA Seq. related research article has been publish in agricultural crops to investigate gene expression and its mechanism. Many RNA-Sequencing platforms have been released in these several years. It also helps to elucidate full-length transcript sequences; as has been demonstrated in a study where ~10% of the untranslated region (*i.e.*, UTR) boundaries of rice genes could be extended [8]. In case of producing genomic and enabling proteomic resources for "non-model" species it is used for example such type of initiative work was done in chickpea [9], safflower [10], pea [11], eucalyptus [12], olive [13], garlic [14] and Japaneese Knotweed [15].

For molecular breeding this tool can be used as De novo in search of genetic polymorphism where thousands of cultivars or very closely related species with variations in traits of interest are sequenced and genetic variation is identified. This allows the generation of molecular markers to facilitate progeny selection and molecular research. Such an illustration of this approaches is the identification of 12,000 single sequence repeats (SSRs) in a single RNA-seq analysis of Sesame, where increased the number of known SSRs from 80 to several thousand with, on average one genic--SSR per ~8 kb [16]. Haseneyer, *et al*. (2011) [17] sampled the transcriptomes of five winter rye inbred lines to identify 5234 Single Nucleotide Polymorphism (SNP). When, they did high-throughput SNP genotyping array and demonstrating the value of RNA sequencing- as a tool for advanced molecular breeding. Characterization of temporal, spatial, regulatory, and evolutionary transcriptome landscapes instances for an example analysis of transcriptomics of grape (Vitis vinifera) berries during three developmental stages identified more than > 6500 genes were expressed in a stage-specific manner [18].

Wang, *et al.* (2012a) [19] analyzed the transcriptome of radish root at two developmental stages and found more than 21,000 genes to be differentially expressed; including genes strongly linking root development first with starch and sucrose metabolism and second with phenylpropanoid biosynthesis level. The comparative sequence analysis of the radish RNAseq data and the *Brassica rapa* genome sequence lead to the discovery of 14,641 SSRs.

The research of Matas, *et al*. (2011) [20] used LCM in-combination with RNA-seq (454 pyrosequencing) for transcriptome profiling of 5 types of tissue in tomato fruit pericarp development. They have found approximately 21,000 unique gene; out of which more than half showed ubiquitous expression whereas other subgroups displayed clear cut cell type-specific expression; these results provides insights into numerous aspects of fruit biology. Stelpflug, *et al*. (2016) [21] studied maize crop for exploring stages of root development. In Maize comprehensive and systematic transcriptome profiling provided valuable insight into root transcriptome which represented 28,894 (73.2%) annotated genes where, they noted transformation across both longitudinally as well as radial gradients of the primary root supported by 4 fold differential expression of 9353 (4 zone) and 4728 (cortical parenchyma and stele) genes. Similarly, Takacs, *et al*. (2012) [22] examined some genes of Maize shoot apical meristem by using RNA-seq coupled with Illumina-based NGS. More than fifty unknown drought responsive genes were discovered by Dugas, *et al*., (2011)[23].

RNA sequencing was done to reveal massive changes in metabolism and cellular physiology of Chlamydomonas reinhardtii green alga under the condition of cellular sulfur deprivation [24]. The most valuable oilseed Brassica napus is under constant pressure of fungal pathogen (*Sclerotinia sclerotiorum*) of white stem rot; by using global RNA sequencing. Girard, *et al*. (2017) [25] profiled gene activity at the first point of infection on the leaf surface 24 hrs after pathogen exposure in susceptible- *B. napus* cv. Westar and tolerant- *B. napus* cv. Zhongyou 821. They have identified a family of ethylene response factors which play role to host tolerance to *S. sclerotiorum*. Physiological investigation of redox homeostasis was then examined by quantifying cellular levels of the glutathione and ascorbate redox pathway and the cycling enzymes associated with host tolerance to *S. sclerotiorum* [25]. A comparative transcriptomic analysis was performed between hot water treated and non-treated fruit of tomato before and after cold storage. In this analysis massive sequencing was done on a 5500 SOLID System with Exact Call Chemistrymodule and identified large number of differentially expressed gene. Three clusters of genes were identified after two weeks of cold storage: the chilling-response included down-regulation of genes involved in photo-synthesis, metabolism of cell wall, lipid and ethylene, as well as the up-regulation of genes for trehalose synthesis and transcription factors DOF and MYB. In their study, down-regulation of genes were involved in carotenoid biosynthesis, which correlates with the main CI symptom of uneven ripening; meanwhile, the chilling-tolerance was interconnected to the up-regulation of genes for heat stress *i.e*., heat shock proteins and heat shock transcription factors and detoxification (glutathione S-transferases) [26].

Liu, *et al*. (2013) [27] reported that Solexa sequencing can be utilized to discover small RNA populations and to compare miRNAs on genome-wide scale in watermelon grafting system. They obtained total of 11.458.476, 11.614.094 and 9.339.089 raw reads representing 2.957.751, 2.880.328 and 2.964.990 unique sequences from the scions of self-grafted watermelon and watermelon grafted on-to bottle gourd and squash at two true-leaf stage, respectively. From small RNA dataset they have identified 39 known miRNAs belonging to 30 miRNA families and 80 novel miRNAs. When watermelon was grafted onto different rootstocks then miRNAs expressed differentially which suggested importance of miRNAs role in diverse biological and metabolic processes in watermelon. Grafting may consequently change miRNAs expressions to regulate stress adaptation, plant growth and development.

Isolation of mRNA

The procedure of isolation of high quality-mRNA from RNA pools takes precedence over the subsequent downstream cascade of RNA sequencing. The relative population of mRNA is only about ~1-5 percent for which special selection approaches are required. Because mature protein coding mRNA contains a poly-A tail, polyadenylated RNA selection is arguably the most prevalent use. To select poly-A + mRNA magnetic or cellulose beads coated with oligo dT molecules can be used. Polyadenylated RNAs can also be chosen for reverse transcription utilizing oligo-dT priming. Depletion of rRNA using sequence-specific probes can be hybridize to rRNAs' which is alternative method of mRNA enrichment. For depletion of unwanted rRNAs or their cDNAs biotinylated DNA or locked nucleic acid (LNA) probes are used to hybridize them and then streptavidin beads are used to deplete them. Antisense DNA oligoscan also be used to target rRNAs, which is known as probe-directed degradation (PDD) [28] although this is less time-consuming than hybridization, but this method necessitates constant rRNA coverage and distinct probe sets for each species.

Preparation of Library

The very crucial step of RNA sequencing is the generation of the library since it influences how closely the cDNA sequence data reflects the original RNA population. The easiest method is to make double stranded cDNA and ligate the adaptor to it. In order to make high quality cDNA it is essential to start with a population of intact mRNAs. Several hundred Adenine bases (AAA…) are present at the 3' terminus of most eukaryotic mRNAs. These poly-A tail can trap these RNAs while also removing contaminated rRNAs, tRNAs, and other small cytoplasmic and as well as nuclear RNAs. For production of DNA copies from mRNA reverse transcriptase can be used with an oligo dT primer [29]. A specific mRNA or class of mRNAs' may be identified by using random primers. As per Stahl, *et al*. (2011) [30] there are two common methods of RNA-Seq experimental protocols: I) single end and II) paired end sequencing procedures. Nucleotide molecules of 50 to 100 base pair (bp) in length or 200 to 400 bp in length can be sequenced from one or both the ends. Method for production of cDNA library preparation for next generation sequencing [31].

Amplification

Amplification is done for amplifying cDNA library to get multiple reads based on the type of RNA sequence used. It must be done before using PCR as most sequencers have a limited detection capacity. Most NGS systems demand library amplification before sequencing except for Pacific Biosciences' (PacBio) single molecule sequencer. This provides quantitative results of RNA expression; however, quantification of expression levels can be due to PCR as well as increased RNA expression [30]. Thus, comparison between multiple samples must be available in order to determine the special effects of PCR on measured expression levels relative to the actual levels in the plant cell. Helicos sequencing does not require amplification of cDNA; and RNA sequencing methods are improving to diminish PCR effects. However, it's very crucial to inspect if PCR amplification will have an impact on the experimental construct [1].

Sequencing by synthesis step

The final step is to identify each of the individual nucleotide that will be used in the sequences. Illumina uses sequencing by synthesis approach that employs reversible terminators in which the four modified nucleotides, sequencing primers and DNA polymerases are added as a mix, and then the primers are hybridized to the sequences. After that, the primers are extended with polymerases by means of the modified nucleotides.

To discriminate each nucleotide type, each is labeled with a fluorescent specific. Because the nucleotides have an inactive 3'-hydroxyl group, only one nucleotide is incorporated [32]. A laser excites clusters, causing them to emit a light signal unique to each nucleotide; which is detected by a coupled charge device CCD camera and translated into a nucleotide sequence by computer programs. The process is repetitive by removing the terminator with the fluorescent label and beginning a new cycle with a new incorporation until the required size reads are synthesized [33].

RNA- sequencing data analysis step

The preliminary step in data processing is to map the short reads from RNA-Sequencing to the reference genome; or to assemble them into contigs before aligning them to the genomic sequence to show transcription structure [34].

The most common goal of this method of transcriptome analysis is to estimate expression of specific genomic areas, which could include genes, isoforms, exons, splice junctions, or newly transcribed regions. It is critical to map the reads obtained from synthesized sequences in order to accomplish this goals.

Read mapping

The proof of identity of which features are present in the sequencing library is the initial stage in read mapping. Because mapping these short reads to the genome is challenging [35]. Total three methods have been used: a. de novo assembly of reads, b. read alignment to the genome, and c. then assembly and read alignment to the transcriptome. Because there is no single method for read alignment; the choice of aligners will be influenced by the reference (genome or transcriptome) utilized, the data type (short vs. longer reads), and the computational capacity available. Short read alignment is currently performed by a variety of techniques, and the preferred aligner is generally determined by the analytic goals and needs [36].

De novo assembly

Plant genomes are having overlaps and these overlaps are utilize for de novo read assembly. It is a method to find a set of the longest possible contiguous expressed regions (contings) [35]. In current time, three algorithmic solutions have been used to overcome the challenge of de novo assembly: prefix tree, overlap-layout-consensus, and de Bruijn graph [36]. Even though de novo read assembly is the most difficult of the three mapping procedures, it is the method of choice when a reference genome is not available or the annotation for the species in question is of poor quality [37].

Read alignment to reference gene or transcriptome

The readings can be mapped to either a genome or a transcriptome as a reference (*i.e.*, the set of all known transcript RNA sequences for a species). Read alignment to the reference genome, on the other hand, gives the benefit of permitting the discovery of new genes and isoforms. Alternative splicing happens concurrently with transcription and is mostly controlled by splicing factors (proteins) that bind RNA motifs (short lengths of RNA) in the pre-mRNA [3]. The problem is solved by means of read mapping tools that can detect alternative splices from the original transcript. TopHat [38], QPALMA and SOAPSplice are some of the alignment tools, which can align spliced junctions of short reads [39].

Differential expression

Read counts in genes/exons are used to determine gene expression and the expression of genes in unlike samples is compared by using gene expression analysis. The alignment result from RNA-sequencing gives the chromosome/position of each aligned read. There are reads aligned to the gene body for each gene which may concise into a number, for the expression by counting the number of reads aligned and normalizing by the total number of reads in the experiment, as well as gene lengths, if desired [28]. The fragments per kilobase of transcript per million fragments mapped (FPKM) technique is used to quantify expression [40]. The number of reads mapped to a gene decides the amount of expression and the amount of reads obtained from an expressed gene is proportional to the length of the transcript. In order to examine means and variances, differential expression normally requires abundant replicates per sample. Normalization for Library size is essential when comparing gene expression between samples group (number of reads obtained). As normalization for gene length and library size is performed by transforming counts to fragments per kilo base per million mapped reads (FPKM), thus the longer the transcript the more reads in the library [41].

Alternative Splicing

Alternative splicing is a transcript processing method by connecting distinct exons together to produce distinct mRNA products from the same pre-mRNA molecule. It has been reported that RNA sequencing was used for the first time to quantify splicing by using a method similar to splicing junction arrays [42]. Exon arrays, on which the probes are tailored to target the junction areas, can be apply to detect and quantify it. At the junctions of two exons, RNA-Sequencing detects "junction reads," which are reads that overlapping area of two exons [43]. Reads that do not align to the genome but map to these synthetic pieces is evidence for splice junctions between recognized exons. Splice sites can be identified from the start by looking for reads that bridges exon junctions; however matching short read to the genome is computationally difficult. Concatenating known neighboring exons and then synthesizing sequence fragments from those spliced transcripts solves this challenge [44, 45]. TopHat, on the other hand, is a software program that uses large-scale mapping of RNA sequencing reads to identify splice sites.

Gene fusing

It is the method in which two or more genes can be “fused” to generate a new gene. Microarrays are ineffective in detecting this phenomenon. Reads from “paired-end” RNA-Seq, on the other hand, provide a wealth of information. In this as the DNA segments are selected based on the sizes, and there should not be long cDNA segments as discussed by Ozsolak and Milos (2011) [46].

Conclusion

Until Watson and Crick discovery of DNA double helix structure (1953) there was a gap of genome knowledge. Later on, this information was utilized for discovery of the human genome project. Before the advent of next-generation sequencing technologies vast understanding about genome structure and function was a challenging task. The advancement of DNA and cDNA sequencing technologies from the transcriptome has helped in gene analysis of various organisms. Major advances in sequencing generation evolution have resulted in the emergence of the most recent RNA sequencing technology. Beyond differential gene expression, this has resulted in a noteworthy paradigm shift in the area of transcriptomics. It enabled scientists to comprehend and analyze various forms of gene expression, such as alternative splicing, isoforms, and fusions in a limited time. The major challenges are mapping the NGS ‘read' to the reference transcriptome as a result, RNA-Sequencing analysis must map to the ‘reference genome’ as a transcriptome proxy. The ambiguous processes in preparation and protocols are the technology's limits. However, bioinformatics and computational components for converting biological data into information require further development in this area.

Application of research: Study of understanding gene expression and different cellular mechanism of plant organ or tool for OMICS research.

Research Category: Plant Breeding and Biotechnology

Acknowledgement / Funding: Authors are thankful to Department of Plant Breeding and Genetics, Dr Rajendra Prasad Central Agricultural University, Pusa, 848125, Bihar, India. Authors are also thankful to Department of Agrometeorology, Dr Kalam Agricultural College, Arrabari, Kishanganj, 855107, Bihar Agricultural University, Sabour, 813210, Bihar, India and Department of Genetics and Plant Breeding, SASRD, Nagaland University, Medziphema, 797106, Nagaland, India

****Principal Investigator or Chairperson of research: Dr Mahamaya Banik**
University: Dr Rajendra Prasad Central Agricultural University, Pusa, 848125, Bihar, India
Research project name or number: Review study

Author Contributions: All authors equally contributed

Author statement: All authors read, reviewed, agreed and approved the final manuscript. Note-All authors agreed that- Written informed consent was obtained from all participants prior to publish / enrolment

Study area / Sample Collection: Dr Rajendra Prasad Central Agricultural University, Pusa, 848125, Bihar, India

Cultivar / Variety / Breed name: Grape (*Vitis vinifera*)

Conflict of Interest: None declared

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.
Ethical Committee Approval Number: Nil

References

[1] Ozhan S., Dicle D. and Yildiz A.K. (2017) *American Journal of Plant Biology. Special Issue, Plant Molecular Biology and Biotechnology*, 2(3-1), 1-7.
[2] Weber A.P.M., Weber K.L., Carr K., Wilkerson C. and Ohlrogge J.B. (2007) *Plant Physiol*, 144, 32–42.
[3] Laetitia B.B.M., Zhangjun F., James J.G. and Jocelyn K.C.R. (2013) *Frontiers in Plant Science*, 4(66), 1-10.
[4] Rani B. and Sharma V.K. (2017) *Agricultural Reviews*, 38(4), 271-281.
[5] Wilhelm T. and Landry J.R. (2009) *Methods*, 48, 249–257.
[6] Zhang S., Ding F., He X., Luo C., Huang G. and Hu Y. (2015a) *Molecular Genetics and Genomics*, 290 (1), 365-375.
[7] Million M. and Feyissa T. (2022) *J. Appl. Res Plant Sci*., 3(2), 236-241.
[8] Lu T., Lu G., Fan D., Zhu C., Li W., Zhao Q., etal. (2010) *Genome Res*., 20, 1238– 1249.
[9] Garg R., Patel R.K., Jhanwar S., Priya P., Bhattacharjee A., Yadav G., et al. (2011) *Plant Physiol*., 156, 1661–1678.
[10] Lulin H., Xiao Y., Pei S., Wen T. and Shangqin H.(2012) *PLoSONE* 7,e38653.
[11] Franssen S.U., Shrestha R.P., Brautigam A., Bornberg-Bauer E. and Weber A.P.M. (2011) *BMC Genomics*, 12,227.
[12] Mizrachi E., Hefer C.A., Ranik M., Joubert F., and Myburg A.A. (2010) *BMC Genomics*, 11,681.
[13] Alagna F., D'Agostino N., Torchia L., Servili M., Rao R., Pietrella M., et al. (2009) *BMC Genomics*, 10,399.
[14] Sun X., Zhou S., Meng F. and Liu S. (2012) Plant Cell, 31, 1823–1828.
[15] Hao D.C., Ge G.B., Xiao P.G., Zhang Y.Y. and Yang L. (2011) *PLoS ONE* 6,e21220.
[16] Zhang H., Wei L., Miao H., Zhang T. and Wang C. (2012) *BMC Genomics*, 13, 316.
[17] Haseneyer G., Schmutzer T., Seidel M., Zhou R., Mascher M., Schon C.C., et al. (2011) *BMC Plant. Biol*, 11,131.
[18] Zenoni S., Ferrarini A., Giacomelli E., Xumerle L., Fasoli M., Malerba G., et al. (2010) *Plant Physiol*, 152, 1787–1795.
[19] Wang S., Wang X., He Q., Liu X., Xu W., Li L., et al. (2012a) *Plant Cell Rep*, 31, 1437–1447.
[20] Matas A.J.,Yeats T.H., Buda G. J., Zheng Y., Chatterjee S., Tohge T., et al. (2011) *Plant Cell*, 23, 3893–3910.
[21] Stelpflug S.C., Sekhon R.S., Vaillancourt B., Hirsch C.N., Buell C.R., de Leon N. and Kaeppler S.M. (2016) *The Plant Genome*, 9(1)
[22] Takacs E.M., Li J., Du C., Ponnala L., Janick-Buckner D., Yu J., et al. (2012) *Plant Cell*, 24, 3219–3234.
[23] Dugas D.V., Monaco M.K., Olsen A., Klein R.R., Kumari S., Ware D., et al. (2011) *BMC Genomics*, 12,514.
[24] Gonzalez-Ballester D., Casero D., Cokus S., Pellegrini M., Merchant, S.S. and Grossman A.R. (2010) *Plant Cell*, 22, 2058–2084.
[25] Girard I.J., Tong C., Becker M.G., Mao X., Huang J., Kievit de T., Fernando W.G.D., Liu S., Belmonte M.F. (2017) *Journal of Experimental Botany*, 68(18), 5079–5091.
[26] Cruz-Mendivil J. A., Lopez-Valenzuela C. L., Calderon Vazquez M.O., Reyes-Moreno V.G.C. and Valdez-Ortiz A. (2015) *Postharvest Biology and Technology*, 99, 141–151.
[27] Liu N., Yang J., Guo S., Xu Y. and Zhang M. (2013) *PloS one*, 8 (2), e57359.
[28] Archer S.K., Shirokikh, N.E. and Preiss T. (2015) *Current Protocols in Human Genetics*, 85(1), 11-15.
[29] Gong A.D., Lian S.B., Wu N.N., Zhou Y.J., Zhao S.Q., Zhang L.M. and Yuan H.Y. (2020) *BMC plant biology*, 20(1), 1-14.
[30] Stahl F., Hitzmann B., Mutz K., Landgrebe D., Lubbecke M., Kasper C., and Scheper T. (2011) *Genomics and Systems Biology of Mammalian Cell Culture*, 1-25.
[31] Wang L., Li P. and Brutnell T.P. (2010b) *Briefings in functional genomics*, 9(2), 118-128.
[32] Kchouk M., Gibrat J.F., and Elloumi M. (2017) *Biology and Medicine*, 9(3).
[33] Backiyarani S., Uma S., Saraswathi M.S., Saravanakumar A.S. and Chandrasekar A. (2015) *Turkish Journal of Agriculture and Forestry*, 39 (5), 705-717.
[34] Hrdlickova R., Toloue M. and Tian B. (2017) *Wiley Interdisciplinary Reviews, RNA*, 8(1), e1364.
[35] Peri S, Roberts S, Kreko IR, McHan LB, Naron A, Ram A, Murphy RL, Lyons E, Gregory BD, Devisetty UK and Nelson ADL. (2020) *Front. Genet,* 10, 1361.
[36] Pandit A.A., Shah R.A. and Husaini A.M. (2018) *J. Pharmac Phytochem*, 7, 1701-1704.
[37] Dida F., & Yi G. (2021) *Peer J Comput Sci,* 7, e636
[38] Trapnell C., Pachter L. and Salzberg S.L. (2009) *Bioinformatics*, 25(9), 1105-1111.
[39] Nagalakshmi U., Waern K. and Snyder M. (2010) *Current protocols in molecular biology*, 89(1), 4-11.
[40] Mortazavi A., Williams B.A., and McCue K. Schaffe er L. and Wold B. (2008) *Nature methods*, 5(7), 621-628.
[41] Smid M., Coebergh van den Braak R.R., van de Werken H.J., van Riet J., van Galen A., de Weerd V. and Sieuwerts A.M. (2018) *BMC Bioinformatics*, 19(1), 1-13.
[42] Alamancos G.P., Agirre E. and Eyras E. (2012) *Methods Mol Biol*, 9, 1-31.
[43] Wang Z., Gerstein M. and Snyder M. (2009, c) *Nature reviews genetics*, 10(1), 57-63.
[44] Orton R.J., Gu Q., Hughes J., Maabar M., Modha S., Vattipally S. and Davison A. (2016) *International Office of Epizootics*, 35(1), 241- 285.
[45] Trapnell C., Hendrickson D.G., Sauvageau M., Goff L., Rinn J.L., Pachter L. (2013) *Nat. Biotechnol*. 31(1), 46-53.
[46] Ozsolak F. and Milos P.M. (2011) *Nature reviews genetics*, 12(2), 87-98.