Journal Homepage: - www.journalijar.com
INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)
Article DOI: 10.21474/IJAR01/3855
DOI URL: http://dx.doi.org/10.21474/IJAR01/3855
ISSN NO. 2320-5407

REVIEW ARTICLE

RNA SEQUENCING AND EUKARYOTIC ORGANISMS, WITH SPECIAL REFERENCE TO *HAEMONCHUS CONTORTUS*: A MINIREVIEW.

Saeed El-Ashram[1,2], Ibrahim Al Nasr[3,4], Fathi Abouhajer[5,6], Rashid mehmood[7,8], Min Hu[9] and *Xun Suo[1].

1. National Animal Protozoa Laboratory, College of Veterinary Medicine, China Agricultural University, Beijing 100193, China.
2. Faculty of Science, Kafr El-Sheikh University, Kafr El-Sheikh, Egypt.
3. College of Science and Arts in Unaizah, Qassim University, Unaizah, Saudi Arabia.
4. College of Applied Health Sciences in Ar Rass, Qassim University, Ar Rass 51921, Saudi Arabia.
5. Asmarya University for Islamic Sciences, Zliten, Misrata, Libya.
6. College of Animal Sciences and Technology, China Agricultural University (CAU), Beijing 100193, China.
7. College of information science and technology, Beijing normal university, Beijing, china.
8. Department of Computer Science and Information Technology, University of Management Sciences and Information Technology, Kotli Azad Kashmir, 11100, Pakistan
9. State Key Laboratory of Agricultural Microbiology, Key Laboratory of Development of Veterinary Products, Ministry of Agriculture, College of Veterinary Medicine, Huazhong Agricultural University, Wuhan 430070, Hubei,China.

Manuscript Info

Manuscript History

Received: 03 February 2017
Final Accepted: 04 March 2017
Published: April 2017

Key words:-
RNA-seq; Haemonchus contortus; ovine host

Abstract

Next-generation sequencing (NGS) technologies are now being employed to disclose the continually changing in the transcriptional profile of a biological sample in an approach called RNA sequencing (RNA-seq). Here, we review how gene expression technology is shifting from conventional to contemporary sequencing approaches. We also provide a snap shot about the application of RNA-seq technology for discovery and comparison of gene expression patterns in different organisms with a particular emphasis on the impact of *Haemonchus contortus* on ovine hosts. Furthermore, we presented steps in conducting RNA-seq experiment and how to validate the RNA-seq data. A future study investigating of *Haemonchus* and its ovine host using RNA-seq technology would be very interesting.

The need for RNA sequencing (RNA-seq):-

Several model organism genomes have been sequenced; the whole genome expression profile under the impact of a particular factor should be investigated for interpreting the genome functional elements, and revealing the cell and tissue molecular constituents, and also for better understanding development and disease. Many methods have been applicable to explore the transcriptome, but only of a few genes at a time. Large-scale screenings of gene expression profiles were not probable the way they have been recently performed with RNA sequencing (RNA-seq). There is, therefore, a definite need for an immediate snapshot of all or a large set of genes. Furthermore, the necessity of RNA-seq emergence is to understand the bio-molecular interaction networks at a global scale. Each specific kind of cell or tissue will be distinguished by a diverse gene expression pattern (i.e. each cell or tissue type will produce a

Corresponding Author:- Xun Suo.
Address:- State Key Laboratory for Agrobiotechnology, College of Veterinary Medicine, China Agricultural University, Beijing 100193, China.

specific set of proteins in very peculiar quantities). Gene inactivation (knocking it out), which was for a long time the only common approach in genetics available to study the impacts of this knockout in other genes, is extremely slow, expensive and inefficient for a large-scale screening of several genes. RNA-seq allows the screening of many genes at the same time. By taking a snapshot of a whole gene expression pattern in a given cell or tissue, many tissues can be compared with each other or a tumor with the healthy tissue surrounding it. Furthermore, the impacts of drugs or stressors on different tissues can be monitored by the gene expression levels. The phenomena related to aging or fetal development can be understood by gene expression. Screening tests can be designed for a myriad of conditions distinguished by specific gene expression patterns. Drug development, diagnosis, comparative genomics, functional genomics and many other fields may benefit hugely from RNA-seq technology that allows accurate and relatively economical collection of gene expression information for many genes at a time. The impact of parasitic elements (e.g. larval-stages, cercarial stages, sporozoites, and tachyzoites) on the different predilection site in the host can be investigated by RNA-seq technology. Gene expression of parasites at different developmental stages can be carried out by NGS technology for discovery and comparison of gene expression patterns. A wide variety of RNA-seq applications have been reviewed in detail elsewhere [1].

RNA sequencing-tool for gene expression analysis:-

The development of molecular approaches, for example, candidate gene method (e.g. real-time polymerase chain reaction) and exploratory technique (e.g. microarrays) has fascilitated the exploration of the gene expression or transcriptional profiles. The current gold standard for protein-coding gene annotation is expressed sequence tag (EST) or full-length cDNA sequencing followed by alignment to a reference genome, but it has been calculated that most EST studies using Sanger sequencing discover approximately 60% of cell transcripts, which declines to disclose the low abundant and long transcripts [2]. This information gap can be addressed by exploiting the RNA-seq technologies. RNA-seq is a vigorous tool to unravel the complex landscape and dynamics of transcriptomes at an exceptional level of sensitivity and accuracy [3; 4; 5]. This approach offers a number of advantages compared to other technologies, including microarrays. These are: unbiased detection of novel transcripts, broader dynamic range, compatible with any species, easier detection of rare and low-abundance transcripts, better estimate of relative expression levels of any genomic region with higher technical reproducibility, facilitating the alternative splicing detection [6; 7]. Along with these advantages, RNA-seq has been employed to reassemble the whole organism transcriptome [8; 9]. With today's advances in RNA-seq technology, enormous sets of gene expression data can be generated. Such catalogues are known as gene expression or transcriptional profiles, and the data collecting process is named profiling. RNA-seq technology allows rapid profiling and deep mining of the transcriptome. While the mRNA-seq application requires especial lab methods (poly-A selection for mRNA purification from total-RNA, reverse transcription into cDNAs), the instrumental rationale for mRNA-seq is similar to that of Genome-seq. As for reference-based mRNA-seq application, illumina single-end or paired-end layouts are favored [10; 11](Table 1).

Table 1:- Advantages of RNA-seq compared with other transcriptomics methods [4].

Technology	Tiling microarray	cDNA or EST sequencing	RNA-seq
Technology specifications			
Principle	Hybridization	Sanger sequencing	High-throughput sequencing
Resolution	From several to 100 bp	Single base	Single base
Throughput	High	Low	High
Reliance on genomic sequence	Yes	No	In some cases
Background noise	High	Low	Low
Application			
Simultaneously map transcribed regions and gene expression	Yes	Limited for gene expression	Yes
Dynamic range to quantify gene expression level	Up to a few-hundredfold	Not practical	>8,000-fold
Ability to distinguish different isoforms	Limited	Yes	Yes
Ability to distinguish allelic expression	Limited	Yes	Yes
Practical issues			
Required amount of RNA	High	High	Low
Cost for mapping transcriptomes of large genomes	High	High	Relatively low

Transcriptional profiling:-

RNA-seq is a currently developed method for transcriptome profiling [4]. Investigations using this approach have altered our understanding of the magnitude and complexity of prokaryotic and eukaryotic transcriptomes [12; 13; 14; 15; 16]. To date, next generation sequencing technologies have been employed to create transcriptomes for diverse species and tissues [13; 15; 17]. For example, a study employed the 454 technology to produce 391,157 EST reads from the brain transcriptome of the wasp *Polistes metricus* [18]. The reads were then aligned to the genome sequence and EST resources from the honeybee, *Apis mellifera*, to annotate *P. metricus* transcripts. Strikingly, the study observed wasp EST matches to 39% of the honeybee mRNAs and detected a robust correlation between the expression levels of the corresponding transcripts from the two species. RNA-Seq has been employed to precisely monitor gene expression during yeast vegetative growth [17], yeast meiosis [19], and mouse embryonic stem (ES) cell differentiation [12], to track gene expression changes during development, and to provide a " digital measurement " of gene expression difference between different tissues. RNA-seq has disclosed diverse novel transcribed regions and splicing isoforms of known genes, and has mapped 5 ′ and 3 ′ boundaries for many genes. In this context, the starts and ends of most transcripts had not been precisely determined, and the extent of spliced heterogeneity remained poorly known before the advent of RNA-seq. Using RNA-seq technology, the 5 ′ and 3 ′ boundaries of 80% and 85% of all annotated genes, respectively, were mapped in *Saccharomyces cerevisiae* [17]. Furthermore, in *Schizosaccharomyces pombe* [19], several boundaries were delineated by RNA-seq data in conjunction with tiling array data. In humans, 31,618 known splicing events were validated (11% of all known splicing events) and 379 novel splicing events were identified [20; 21]. In mice, extensive alternative splicing was charcterized for 3462 genes [13]. Moreover, results from RNA-seq propose the existence of many novel transcribed regions in every genome assessed, including those of *Arabidopsis thaliana* [22], mouse [12; 13], human [20], *S. cerevisiae* [17], and *S. pombe* [19]. The high-throughput paired-end Illumina technology was employed to explore the haemocytes of *O. vulgaris* transcriptome (*de novo* sequencing), identify genes involved in immune defense, and understand the molecular basis of *octopus* tolerance/resistance to coccidiosis [23]. Furthermore, dual RNA-seq of parasite and host reveals gene expression dynamics during filarial worm *Brugia malayi*-mosquito *Aedes aegypti* interactions [24]. The transcriptional profiles of the parasitic nematode *Strongyloides stercoralis* disclose different regulation of canonical dauer pathways [25]. High-throughput RNA sequencing (RNA-seq) has played a crucial in providing a concise view of the *Leishmania major* promastigote stage global transcriptome [9], establishing and enlightening current expression datasets, and providing a solid foundation for drug discovery and vaccine development[26], and studying of the peripheral-blood mononuclear cells (PBMCs) transcriptome from *Fasciola hepatica*-infected sheep[27]. A recent study by [28] examined the transcriptome profiling of differentially expressed genes of *H. contortus*- infected resistant Canaria Hair Breed (CHB) and susceptible Canaria Sheep (CS).

RNA-seq experiment, data generation and analysis:

All RNA-seq experiments follow a similar protocol. The currently used method can be listed as follows:-

RNA extraction:-

Total RNA from fundic abomasal samples of sheep was isolated employing Trizol (Invitrogen, Carlsbad, CA, USA) followed by DNase digestion, as previously reported [29; 30; 31]. 1% agarose gels was exploited to monitor RNA degradation and contamination. The Nano Photometer® spectrophotometer (IMPLEN, CA, USA) was used for checking RNA purity. Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) in combination with its RNA Nano 6000 Assay Kit was used for RNA integrity assessment.

Library preparation:-

RNA-seq library was prepared according to the procedure used by [32] as follows: A total amount of 1 μg RNA per sample was exploited as input material. NEBNext®Ultra™ RNA Library Prep Kit for Illumina®NEB, USA) following manufacturer's recommendations was conducted for the generation of sequencing libraries. For sorting and identification of sequences to each sample, index codes were added. Briefly, the poly (A)-containing mRNA molecules were purified from total RNA employing poly-T oligo-attached magnetic beads. Following purification, the poly (A)-containing mRNA molecules were fragmented using divalent cations in NEBNext First Strand Synthesis Reaction Buffer (5X) under elevated temperature. The first strand cDNA was synthesized from the cleaved RNA fragments using reverse transcriptase and random primers. Subsequently, the second strand cDNA synthesis was carried out exploiting DNA polymerase I and RNase H. T4 DNA polymerase and Klenow DNA polymerase were used to convert overhangs into blunt ends via exonuclease/polymerase activities. Furthermore, NEBNext Adaptor with hairpin loop structure was ligated to prepare for hybridization. Following adaptor ligation,

cDNA fragments of preferentially (approximately 250 to 300 bp) were selected on a gel and the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Adaptor-ligated cDNAs were treated with 3 µl Uracil-Specific Excision Reagent enzyme mix (USER; NEB) at 37 °C for 15 min followed by heat inactivation at 95 °C for 5 min. The clonal amplification of the fragments was conducted with NEBNext Q5 Hot Start HiFi PCR Master Mix, Universal PCR primers and Index (X) Primer, and subsequent purification of PCR products (AMPure XP system) and evaluation of library quality on the Agilent Bioanalyzer 2100 system. For information about the impact of RNA extraction methods and library selection schemes on RNA-seq data, we direct the reader to the published articles [33; 34]

Clustering and sequencing:-
The index-coded samples were clustered on a cBot Cluster Generation System employing the TruSeq PE Cluster Kit v3-cBot-HS (Illumina, San Diego, USA), as stated in the manufacturer's instructions. Then, the library preparations were sequenced with the generation of 150 bp pair-end reads on an Illumina Hiseq 4000 platform. We refer the reader to [35] for the current situation.

Data processing and quality control:-
Raw data (raw reads) in FASTQ format were processed to obtain clean data (clean reads) by trimming the adapter sequences out of the reads (Trimmomatic software v0.33), and filtering read- containing ploy-Ns (Ns>10% in a read), low quality reads (Q<=20) greater than 50% using in-house C scripts. Consequently, the Q20, Q30, GC content of the clean data were calculated. All downstream analyses were carried on clean and high-quality data. The clean reads were aligned to the reference genome using Tophat2 (v2.1.0)[36].

Differential gene expression (DGE) Analysis:-
Differential expression analysis was conducted employing DESeq 2 packages [27] for comparisons among sample gene from different experimental conditions. To determine the statistically significant differential expression, corrected *P*-value (*q*-value) < 0.05 and $|\log_2$ (fold change)| > 1 were set as the threshold for significantly DEGs [37].

Functional annotation enrichment analyses:-
Gene Ontology (GO) database and Kyoto Encyclopedia of Genes and Genomes (KEGG) database were selected to perform DEGs enrichment analysis under different experimental conditions. Goseq R Bioconductor package was implemented for performing Gene Ontology analysis of RNA-seq data [38]. GO terms with adjusted *P*-value less than 0.05 were considered as significantly enriched transcripts. KOBAS v2.0 (KEGG Orthology Based Annotation System) was used to identify the statistical enrichment of differentially expressed genes in KEGG pathways employing hypergeometric test [39]. KEGG terms with corrected *P*-value less than 0.05 were considered statistically significantly enriched genes (Fig. 1) (https://www.illumina.com/techniques/sequencing/rna-sequencing.html). The commonly employed RNA-Seq term explanation is illustrated in Table 2.

Protein-protein interaction networks (PPI):-
Interactions between proteins can be predicted through an array of computational methods and databases [40; 41] .

Novel transcripts prediction:-
Reference Annotation Based Transcript (RABT) assembly method was built upon the Cufflinks v2.1.1 by constructing and identifying both known and novel transcripts from TopHat alignment results in the context of an existing annotation.

Analysis of alternative splicing:-
Alternative splicing events (ASEs) were categorized into 12 basic types, including alternative 5′ first exon (TSS), alternative 3′ last exon (TTS), skipped exon (SKIP), approximate SKIP (XSKIP), multi-exon SKIP (MSKIP), approximate MSKIP (XMSKIP), intron retention (IR), approximate IR (XIR), multi-IR (MIR), approximate MIR (XMIR), alternative exon ends (AE), and approximate AE (XAE), by the software Asprofile v1.0 [42].

Single nucleotide polymorphisms (SNP) Analysis:-
Picard-tools v1.96 (http://sourceforge.net/projects/picard/files/picard-tools/1.96/) and samtools v0.1.18 [43] were employed to classify, remove coupled reads and merge the bam alignment results of each sample. The Genome Analysis Toolkit (GATK2) software was adopted to conduct SNP calling [42; 44].

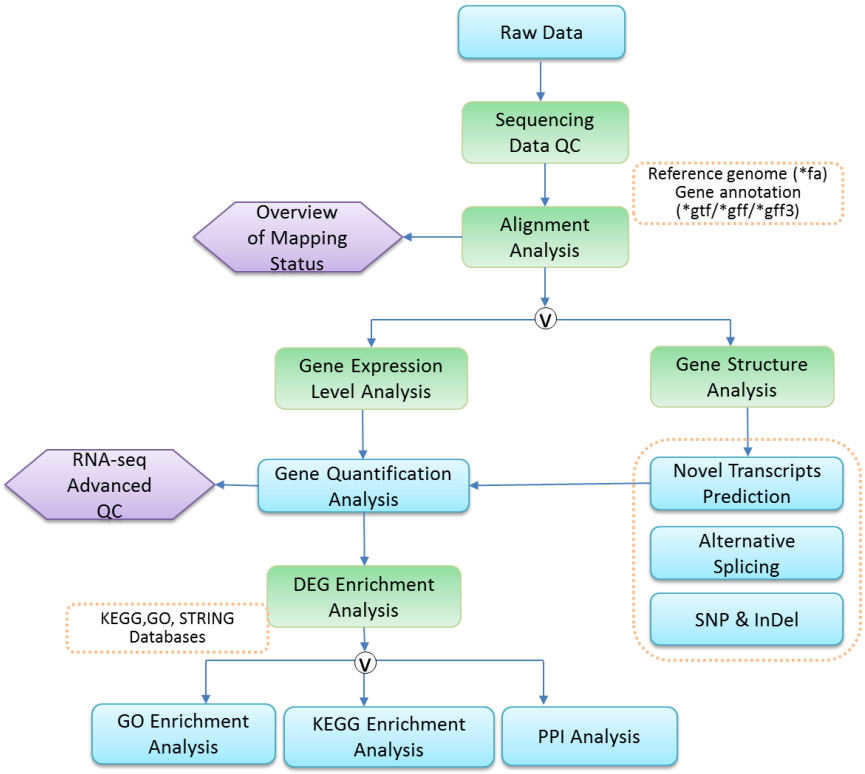

Figure 1:- An RNA-Seq analysis workflow (Beijing Allwegene Technology Co., Ltd, China).

Quantitative reverse transcriptase-PCR (RT-PCR) analysis for validation of RNA-seq Results:-

Since the entire transcript is assessed in a more or less unbiased manner, probe bias, poor sensitivity and reduced linear range are not as problematic in RNA-seq experiments. However, real-time-PCR discrepancies may be due to its oligo (dT) primer and probe-bias based on what region of the cDNA is amplified [4; 45; 46]. A large and growing body of literature has reported a strong correlation between these methods [47; 48; 49; 50; 51; 52; 53]. These results are consistent with our recent data as illustrated below. We conducted transcriptome sequencing of the ovine abomasal tissues using the Illumina Hi Seq 4000 platform to segregate early and late *H. contortus*-infected sheep (7 and 50 days post-infected groups [7 dpi and 50 dpi], respectively) from the control naive ones. We accredit the reader to reviews and articles by [54; 55; 56; 57] for detailed information about *Haemonchus contortus* and its ovine host. By random selection, 13 genes with (overexpressed or repressed) and without differential expression were chosen for verification by quantitative RT-PCR, which was performed as follows: the same total RNA employed for RNA-seq was reverse transcribed employing *EasyScript*® Reverse Transcriptase (Beijing TransGen Biotech Co., Ltd) and SYBR green-based RT-PCR was conducted by using SYBR® Select Master Mix (Applied Biosystems; Cat: 4472908) according to the instructions made by the manufacturer. The results were expressed as fold-changes [58]. A Correlation analysis (Graphpad Software, San Diego, Calif) was performed between the RNA-seq and RT-PCR fold-change results using the same RNA samples before pooling. Additionally, experiments were conducted in triplicate, and data are displayed as mean ± SD.

Validation of transcriptome results by real-time PCR:-

To validate the transcriptome data, both differentially and non-differentially expressed genes were selected for real-time polymerase chain reaction (PCR) analysis, which also showed similar trends concordant with the Illumina sequencing data indicating the reliability of the comparative analysis of our transcriptomes. As expected, transcript-specific fold-change in the same RNA samples was highly consistent between the RNA-seq and RT-PCR methods, which were corroborated in the correlation analysis. For the selected 13 DEGs (Fig. 2A, B, C, D and E), there was a firm correlation between RNA-seq and RT-PCR results ($r^2 = 0.9998$), substantiating the reliability of differential gene expression analysis adopting RNA-seq.

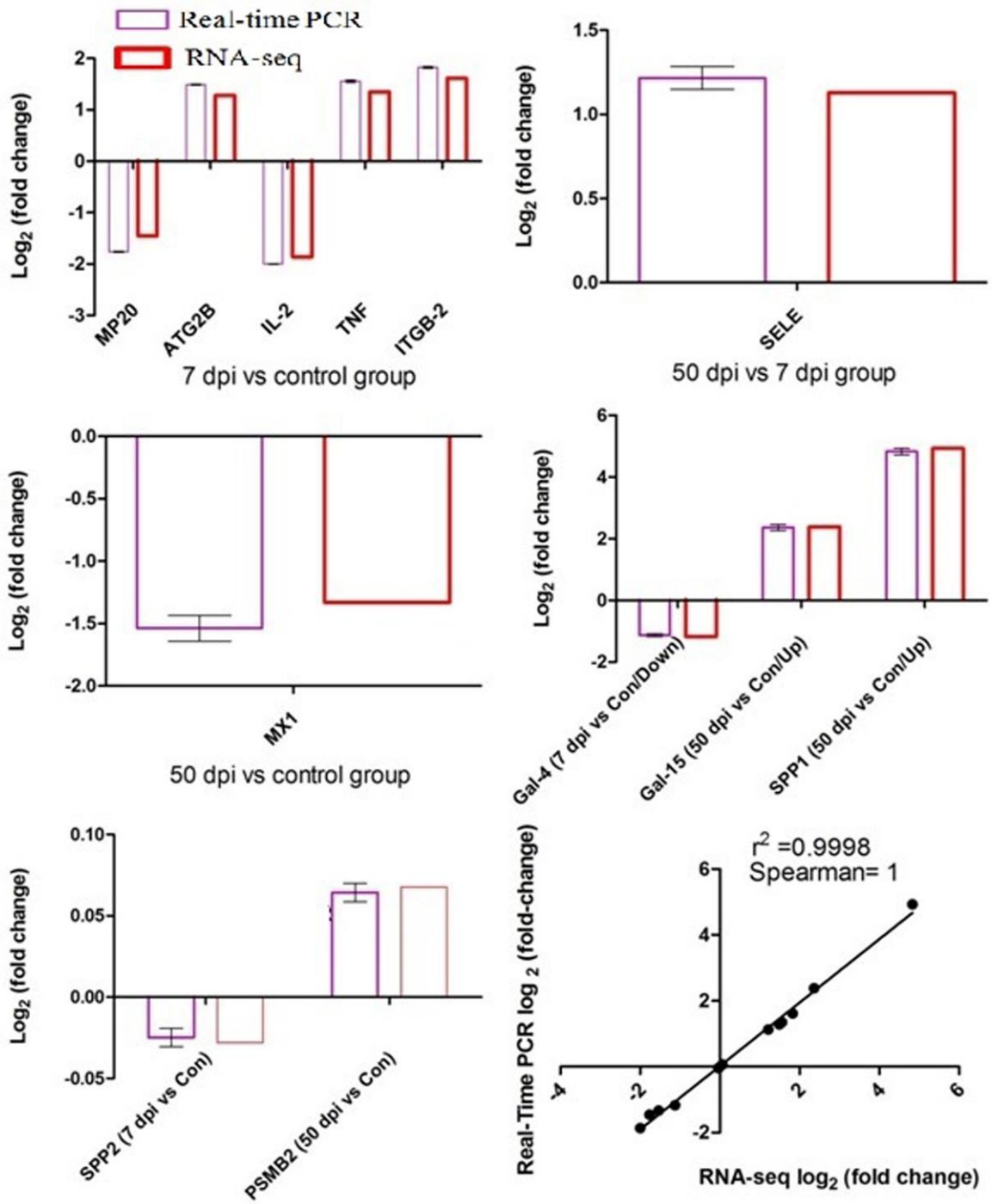

Figure 2:- Validation of RNA-seq data by employing real-time PCR.

The expression of interleukin-2 (IL-2), integrin subunit beta-2 (ITGB-2), tumor necrosis factor (TNF), Autophagy-related 2B (ATG2B), Tumor necrosis factor (TNF), Autophagy related 2B (ATG2B), transcript variant X1, CNDP dipeptidase 2 (metallopeptidase M20 family) (CNDP2), MX dynamin like GTPase-1 (MX-1) and Selectin -E (SELE). Further, the selected expressed (galectin-4 (Gal-4), gaalectin-15 (Gal-15), and SPP1) and non-expressed Proteasome subunit beta 2 (PSMB2) and Secreted phosphoprotein 2 (SPP2) genes are displayed (E). X-axis, gene and group name; Y-axis, log_2 fold change in gene expression (A; B; C; D; E).

F: Correlation analysis between RNA-seq and RT-PCR $\log_2$ (fold-change) results from the same RNA samples. Spearman correlation coefficient is displayed. The coefficient of determination (r^2) expresses the strength of the relationship between the RNA seq and real-time PCR variables.

Table 2:- The commonly employed RNA-seq term explanation.

Term	Explanation	Reference
Single- or Paired-end reads	Single-end reads are enough for regular gene expression profiling, paired-end reads enable discovery applications such as detecting or characterizing novel alternative splicing isoforms or gene fusion events.	[59]
Read length	It relies on the desired results of the experiment. For gene expression profiling, 50 bp single-end reads would be sufficient for most studies.. For detecting currently unknown transcripts, novel splicing isoforms, gene fusion, etc., longer (150 bp) reads offer an advantage.	[60]
Read depth or coverage	Coverage = (total number of bases generated) / (size of genome sequenced). In other words, the average number of reads that align to, or "cover," known reference bases. Low-expressing gene measurement or novel feature identification needs more coverage.	[61]
Complementary DNA (cDNA)	Any DNA that is obtained from an mRNA template via reverse transcription.	[62]
Expressed sequence tags (EST)	cDNA (sub)sequences derived from a single read of a cDNA sequencing experiment.	[62]
RNA-seq	High-throughput sequencing technology utilization to describe entire transcriptomes.	[4]
Transcriptome	The complete set of transcripts in a cell, and their quantity, for a specific developmental stage or physiological condition.	[63; 64]
Model	Genome fully sequenced and annotated	[65]
Novel	little/no previous sequencing	[65]
Barcodes (index)	Short (6-8 nt) introduced as part of adapters. It provides unique identifier for each sample, tolerance of 1-2 sequencing errors, pooling samples to mitigate lane effects, and allowing deep multiplexing due to dual barcodes.	[66]

Conclusion and future directions:-

We have attempted to provide a snap shot of RNA-seq technology as a tool for gene expression analysis. Future research may exploit RNA-seq to provide a dual RNA-seq time course analysis of *H. contortus* and ovine host.

Acknowledgments:-

This research was supported by the National Key Basic Research Program (973 program) of China (Grant No. 2015CB150300). Additionally, the funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests:-

The authors declare that they have no competing interests.

References:-

1. A. Conesa, P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M.W. Szczesniak, D.J. Gaffney, L.L. Elo, X. Zhang, and A. Mortazavi, A survey of best practices for RNA-seq data analysis. Genome Biol 17 (2016) 13.
2. M.R. Brent, Steady progress and recent breakthroughs in the accuracy of automated genome annotation. Nature reviews. Genetics 9 (2008) 62-73.
3. F. Ozsolak, and P.M. Milos, RNA sequencing: advances, challenges and opportunities. Nature reviews. Genetics 12 (2011) 87-98.
4. Z. Wang, M. Gerstein, and M. Snyder, RNA-Seq: a revolutionary tool for transcriptomics. Nature reviews. Genetics 10 (2009) 57-63.
5. S. Marguerat, and J. Bahler, RNA-seq: from technology to biology. Cellular and molecular life sciences : CMLS 67 (2010) 569-79.
6. Y. Yao, Z. Ni, Y. Zhang, Y. Chen, Y. Ding, Z. Han, Z. Liu, and Q. Sun, Identification of differentially expressed genes in leaf and root between wheat hybrid and its parental inbreds using PCR-based cDNA subtraction. Plant Mol Biol 58 (2005) 367-84.
7. J.C. Marioni, C.E. Mason, S.M. Mane, M. Stephens, and Y. Gilad, RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. Genome Res 18 (2008) 1509-17.
8. T.N. Siegel, K. Gunasekera, G.A. Cross, and T. Ochsenreiter, Gene expression in Trypanosoma brucei: lessons from high-throughput RNA sequencing. Trends Parasitol 27 (2011) 434-41.
9. A. Rastrojo, F. Carrasco-Ramiro, D. Martin, A. Crespillo, R.M. Reguera, B. Aguado, and J.M. Requena, The transcriptome of Leishmania major in the axenic promastigote stage: transcript annotation and relative expression levels by RNA-seq. BMC Genomics 14 (2013) 223.
10. C. Tachibana, Transcriptomics today: Microarrays, RNA-seq, and more. Science Science 349 (2015) 544-546.
11. S. Zhao, W.-P. Fung-Leung, A. Bittner, K. Ngo, X. Liu, and S.-D. Zhang, Comparison of RNA-Seq and Microarray in Transcriptome Profiling of Activated T Cells. PLoS ONE PLoS ONE 9 (2014) e78644.
12. N. Cloonan, A.R. Forrest, G. Kolle, B.B. Gardiner, G.J. Faulkner, M.K. Brown, D.F. Taylor, A.L. Steptoe, S. Wani, G. Bethel, A.J. Robertson, A.C. Perkins, S.J. Bruce, C.C. Lee, S.S. Ranade, H.E. Peckham, J.M. Manning, K.J. McKernan, and S.M. Grimmond, Stem cell transcriptome profiling via massive-scale mRNA sequencing. Nat Methods 5 (2008) 613-9.
13. A. Mortazavi, B.A. Williams, K. McCue, L. Schaeffer, and B. Wold, Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5 (2008) 621-8.
14. D.J. Sugarbaker, W.G. Richards, G.J. Gordon, L. Dong, A. De Rienzo, G. Maulik, J.N. Glickman, L.R. Chirieac, M.L. Hartman, B.E. Taillon, L. Du, P. Bouffard, S.F. Kingsmore, N.A. Miller, A.D. Farmer, R.V. Jensen, S.R. Gullans, and R. Bueno, Transcriptome sequencing of malignant pleural mesothelioma tumors. Proceedings of the National Academy of Sciences of the United States of America 105 (2008) 3521-6.
15. M. Sultan, M.H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, D. Schmidt, S. O'Keeffe, S. Haas, M. Vingron, H. Lehrach, and M.L. Yaspo, A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science (New York, N.Y.) 321 (2008) 956-60.
16. F. Tang, C. Barbacioru, E. Nordman, B. Li, N. Xu, V.I. Bashkirov, K. Lao, and M.A. Surani, RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nature protocols 5 (2010) 516-35.
17. U. Nagalakshmi, Z. Wang, K. Waern, C. Shou, D. Raha, M. Gerstein, and M. Snyder, The transcriptional landscape of the yeast genome defined by RNA sequencing. Science (New York, N.Y.) 320 (2008) 1344-9.
18. A.L. Toth, K. Varala, T.C. Newman, F.E. Miguez, S.K. Hutchison, D.A. Willoughby, J.F. Simons, M. Egholm, J.H. Hunt, M.E. Hudson, and G.E. Robinson, Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. Science (New York, N.Y.) 318 (2007) 441-4.
19. B.T. Wilhelm, S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C.J. Penkett, J. Rogers, and J. Bahler, Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453 (2008) 1239-43.
20. R. Morin, M. Bainbridge, A. Fejes, M. Hirst, M. Krzywinski, T. Pugh, H. McDonald, R. Varhol, S. Jones, and M. Marra, Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. BioTechniques 45 (2008) 81-94.
21. R.D. Morin, M.D. O'Connor, M. Griffith, F. Kuchenbauer, A. Delaney, A.L. Prabhu, Y. Zhao, H. McDonald, T. Zeng, M. Hirst, C.J. Eaves, and M.A. Marra, Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. Genome Res 18 (2008) 610-21.

22. R. Lister, R.C. O'Malley, J. Tonti-Filippini, B.D. Gregory, C.C. Berry, A.H. Millar, and J.R. Ecker, Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell 133 (2008) 523-36.
23. S. Castellanos-Martinez, C. Gestal, D. Arteta, and S. Catarino, De novo transcriptome sequencing of the octopus vulgaris hemocytes using illumina RNA-Seq technology: Response to the infection by the gastrointestinal parasite Aggregata octopiana. PLoS ONE PLoS ONE 9 (2014).
24. Y.J. Choi, M.T. Aliota, G.F. Mayhew, S.M. Erickson, and B.M. Christensen, Dual RNA-seq of parasite and host reveals gene expression dynamics during filarial worm-mosquito interactions. PLoS neglected tropical diseases 8 (2014).
25. J.D. Stoltzfus, S. Minot, M. Berriman, T.J. Nolan, and J.B. Lok, RNAseq Analysis of the Parasitic Nematode Strongyloides stercoralis Reveals Divergent Regulation of Canonical Dauer Pathways. PLoS neglected tropical diseases 6 (2012) e1854.
26. L.A. Dillon, K. Okrah, V.K. Hughitt, R. Suresh, Y. Li, M.C. Fernandes, A.T. Belew, H. Corrada Bravo, D.M. Mosser, and N.M. El-Sayed, Transcriptomic profiling of gene expression and RNA processing during Leishmania major differentiation. Nucleic acids research 43 (2015) 6799-813.
27. C.A. Alvarez Rojas, J.P. Scheerlinck, B.R. Ansell, R.S. Hall, R.B. Gasser, and A.R. Jex, Time-Course Study of the Transcriptome of Peripheral Blood Mononuclear Cells (PBMCs) from Sheep Infected with Fasciola hepatica. PLoS One 11 (2016) e0159194.
28. Z. Guo, J.F. Gonzalez, J.N. Hernandez, T.N. McNeilly, Y. Corripio-Miyar, D. Frew, T. Morrison, P. Yu, and R.W. Li, Possible mechanisms of host resistance to Haemonchus contortus infection in sheep breeds native to the Canary Islands. Scientific reports 6 (2016) 26200.
29. R.W. Li, and L.C. Gasbarre, A temporal shift in regulatory networks and pathways in the bovine small intestine during Cooperia oncophora infection. Int J Parasitol 39 (2009) 813-24.
30. J. Lv, P. Liu, Y. Wang, B. Gao, P. Chen, and J. Li, Transcriptome analysis of Portunus trituberculatus in response to salinity stress provides insights into the molecular basis of osmoregulation. PLoS One 8 (2013) e82155.
31. S. El-Ashram, I. Al Nasr, and X. Suo, Nucleic acid protocols: Extraction and optimization. Biotechnology reports (Amsterdam, Netherlands) 12 (2016) 33-39.
32. X. Li, X. Zuo, J. Jing, Y. Ma, J. Wang, D. Liu, J. Zhu, X. Du, L. Xiong, Y. Du, J. Xu, X. Xiao, J. Wang, Z. Chai, Y. Zhao, and H. Deng, Small-Molecule-Driven Direct Reprogramming of Mouse Fibroblasts into Functional Neurons. Cell Stem Cell 17 (2015) 195-203.
33. M. Sultan, V. Amstislavskiy, T. Risch, M. Schuette, S. Dokel, M. Ralser, D. Balzereit, H. Lehrach, and M.L. Yaspo, Influence of RNA extraction methods and library selection schemes on RNA-seq data. BMC Genomics 15 (2014) 675.
34. Y. Guo, S. Zhao, Q. Sheng, M. Guo, B. Lehmann, J. Pietenpol, D.C. Samuels, and Y. Shyr, RNAseq by Total RNA Library Identifies Additional RNAs Compared to Poly(A) RNA Library. BioMed research international 2015 (2015) 862130.
35. R. Mehmood, S. El-Ashram, R. Bie, H. Dawood, and A. Kos, Clustering by fast search and merge of local density peaks for gene expression microarray data. Scientific reports 7 (2017) 45602.
36. D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S.L. Salzberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology 14 (2013) R36.
37. Y. Benjamini, and Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 57 (1995) 289-300.
38. M.D. Young, M.J. Wakefield, G.K. Smyth, and A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias. Genome biology 11 (2010) R14.
39. C. Xie, X. Mao, J. Huang, Y. Ding, J. Wu, S. Dong, L. Kong, G. Gao, C.Y. Li, and L. Wei, KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. Nucleic acids research 39 (2011) W316-22.
40. H. Rakshit, N. Rathi, and D. Roy, Construction and Analysis of the Protein-Protein Interaction Networks Based on Gene Expression Profiles of Parkinson's Disease. PloS one 9 (2014) e103047.
41. K. Raman, Construction and analysis of protein-protein interaction networks. Autom Exp 2 (2010) 2.
42. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M.A. DePristo, The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20 (2010) 1297-303.
43. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin, The Sequence Alignment/Map format and SAMtools. Bioinformatics 25 (2009) 2078-9.

44. M.A. DePristo, E. Banks, R. Poplin, K.V. Garimella, J.R. Maguire, C. Hartl, A.A. Philippakis, G. del Angel, M.A. Rivas, M. Hanna, A. McKenna, T.J. Fennell, A.M. Kernytsky, A.Y. Sivachenko, K. Cibulskis, S.B. Gabriel, D. Altshuler, and M.J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43 (2011) 491-8.
45. M. Filion, Quantitative real-time PCR in applied microbiology, Caister Academic Press, Norfolk, UK, 2012.
46. T.R. Hughes, 'Validation' in genome-scale research. Journal of Biology 8 (2009) 3-3.
47. M. Griffith, O.L. Griffith, J. Mwenifumbo, R. Goya, A.S. Morrissy, R.D. Morin, R. Corbett, M.J. Tang, Y.C. Hou, T.J. Pugh, G. Robertson, S. Chittaranjan, A. Ally, J.K. Asano, S.Y. Chan, H.I. Li, H. McDonald, K. Teague, Y. Zhao, T. Zeng, A. Delaney, M. Hirst, G.B. Morin, S.J. Jones, I.T. Tai, and M.A. Marra, Alternative expression analysis by RNA sequencing. Nat Methods 7 (2010) 843-7.
48. Y.W. Asmann, E.W. Klee, E.A. Thompson, E.A. Perez, S. Middha, A.L. Oberg, T.M. Therneau, D.I. Smith, G.A. Poland, E.D. Wieben, and J.P. Kocher, 3' tag digital gene expression profiling of human brain and universal reference RNA using Illumina Genome Analyzer. BMC Genomics 10 (2009) 531.
49. A.R. Wu, N.F. Neff, T. Kalisky, P. Dalerba, B. Treutlein, M.E. Rothenberg, F.M. Mburu, G.L. Mantalas, S. Sim, M.F. Clarke, and S.R. Quake, Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods 11 (2014) 41-6.
50. Y. Shi, and M. He, Differential gene expression identified by RNA-Seq and qPCR in two sizes of pearl oyster (Pinctada fucata). Gene 538 (2014) 313-22.
51. E. Li, S. Wang, C. Li, X. Wang, K. Chen, and L. Chen, Transcriptome sequencing revealed the genes and pathways involved in salinity stress of Chinese mitten crab, Eriocheir sinensis. Physiological Genomics 46 (2014) 177-190.
52. X. Song, X. Hu, B. Sun, Y. Bo, K. Wu, L. Xiao, and C. Gong, A transcriptome analysis focusing on inflammation-related genes of grass carp intestines following infection with Aeromonas hydrophila. Scientific reports 7 (2017) 40777.
53. C. Wang, B. Gong, P.R. Bushel, J. Thierry-Mieg, D. Thierry-Mieg, J. Xu, H. Fang, H. Hong, J. Shen, Z. Su, J. Meehan, X. Li, L. Yang, H. Li, P.P. Labaj, D.P. Kreil, D. Megherbi, S. Gaj, F. Caiment, J. van Delft, J. Kleinjans, A. Scherer, V. Devanarayan, J. Wang, Y. Yang, H.-R. Qian, L.J. Lancashire, M. Bessarabova, Y. Nikolsky, C. Furlanello, M. Chierici, D. Albanese, G. Jurman, S. Riccadonna, M. Filosi, R. Visintainer, K.K. Zhang, J. Li, J.-H. Hsieh, D.L. Svoboda, J.C. Fuscoe, Y. Deng, L. Shi, R.S. Paules, S.S. Auerbach, and W. Tong, The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. Nat Biotech 32 (2014) 926-932.
54. S. El-Ashram, and X. Suo, Corrigendum: Electrical cream separator coupled with vacuum filtration for the purification of eimerian oocysts and trichostrongylid eggs. Scientific reports 7 (2017) 45799.
55. S. El-Ashram, and X. Suo, Exploring the microbial community (microflora) associated with ovine Haemonchus contortus (macroflora) field strains. Scientific reports 7 (2017) 70.
56. S. El-Ashram, and X. Suo, Electrical cream separator coupled with vacuum filtration for the purification of eimerian oocysts and trichostronglyid eggs. Scientific reports 7 (2017) 43346.
57. S. El-Ashram, S. El-Ashram, and X. Suo, Electrical cream separator coupled with vacuum filtration for the purification of eimerian oocysts and trichostrongylid eggs. Protocol Exchange Protocol Exchange (2017).
58. C. Mata-Perez, B. Sanchez-Calvo, J.C. Begara-Morales, F. Luque, J. Jimenez-Ruiz, M.N. Padilla, J. Fierro-Risco, R. Valderrama, A. Fernandez-Ocana, F.J. Corpas, and J.B. Barroso, Transcriptomic profiling of linolenic acid-responsive genes in ROS signaling from RNA-seq data in Arabidopsis. Frontiers in plant science 6 (2015) 122.
59. J.J. Werner, D. Zhou, J.G. Caporaso, R. Knight, and L.T. Angenent, Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. The ISME journal 6 (2012) 1273-1276.
60. S. Chhangawala, G. Rudy, C.E. Mason, and J.A. Rosenfeld, The impact of read length on quantification of differentially expressed genes and splice junction detection. Genome Biology 16 (2015) 131.
61. D. Sims, I. Sudbery, N.E. Ilott, A. Heger, and C.P. Ponting, Sequencing depth and coverage: key considerations in genomic analyses. Nature reviews. Genetics 15 (2014) 121-132.
62. J. Parkinson, and M. Blaxter, Expressed Sequence Tags: An Overview. (2009).
63. B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Global quantification of mammalian gene expression control. Nature 473 (2011) 337-42.
64. B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, and M. Selbach, Corrigendum: Global quantification of mammalian gene expression control. Nature 495 (2013) 126-7.

65. E. Ghedin, S. Wang, J.M. Foster, and B.E. Slatko, First sequenced genome of a parasitic nematode. Trends Parasitol 20 (2004) 151-3.
66. B. LaFlamme, Molecular barcodes improve RNA-seq. Nature genetics 46 (2014) 219-219.