Journal of Next Generation Sequencing & Applications

Ahmad and Mehmood, Next Generat Sequenc & Applic 2015, 2:1
DOI: 10.4172/2469-9853.1000e103

Editorial Open Access

# RNA-Seq: A Powerful Tool for Cataloguing the Transcriptome

**Niaz Ahmad**[1*] **and Muhammad Aamer Mehmood**[2]

*1Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE), Jhang Road, Faisalabad, Pakistan*

*2Department of Bioinformatics & Biotechnology, Government College University Faisalabad, Faisalabad, Pakistan*

***Corresponding author:** Niaz Ahmad, Agricultural Biotechnology Division, National Institute for Biotechnology and Genetic Engineering (NIBGE), Jhang Road, Faisalabad, Pakistan, Tel: +92-419201316-20, Ext 309; Fax: +92(0)41 9201472; E-mail: niazbloch@yahoo.com

**Rec date:** Feb 28, 2015**; Acc date:** May 28, 2015**; Pub date:** May 30, 2015



## Editorial

The Sanger dideoxy chain termination procedure [1] has been a method of choice for determining the sequence of DNA for more than three decades. The method was used to sequence the human genome producing 3GB data over 13 years at a cost of US $ 3.0 billion [2]. This massive amount of information revolutionized the era of genomics. However, it would not be feasible to continue obtaining whole genome sequences at this lengthy pace and cost. The complex questions in the field of biology demand the level of information that is beyond the realms of traditional sequencing methods [3]. The search for alternative methods resulted in "Next Generation Sequencing" or "NGS" technologies [3-5]. Today, we are witnessing the rise and growing applications of NGS: in obtaining information of individual genes from targeted sequencing [4,6], identifying DNA variants that result in disease through exon sequencing [7] and quantifying differentially expressed genes [8], as well as the discovery of new species by metagenomics [9], for example. The high-degree of resolution provided by the NGS techniques not only allows identification of mutations in an organism, it is also helpful in correcting errors in reference genome sequences through resequencing [5].

Although NGS technologies have greatly reduced the cost and speed of obtaining nucleotide sequences, in comparison to the Sanger sequencing, obtaining whole genome sequences remains an expensive and daunting task. An interim solution to this issue is to use NGS platforms for obtaining sequences of targeted regions of the genome; for example, sequencing all the exons or particular gene families associated with a particular phenotype, condition, or disease. One such approach is to catalogue the transcriptome by RNA sequencing (RNA-Seq) to obtain an insight into the various functional elements of the genome involved in organismal development or those associated with a particular set of physical conditions [4,8]. Earlier methods of analyzing the transcriptome used hybridization-based techniques that required genome sequence information, produced high backgrounds, and had low detection limits [10]. Due to these factors, hybridization-based approaches are difficult to apply to comparative gene expression studies, as they often need complicated and intensive normalization. In contrast to hybridization-based methods, sequencing methods rely on cDNA sequence. Early implementations used Sanger sequencing of cDNA libraries or EST libraries and were relatively low throughput, and expensive. RNA-Seq has become popular in the scientific community over hybridization (microarray)-based gene expression studies since it provides a comprehensive analysis of gene expression down to a single base resolution [8], thereby identifying overlapping transcripts, novel transcripts, as well as providing information on transcription start/endsites [11]. RNA-Seq has become a powerful tool to obtain transcript information for those organisms for which genome sequences are not available. RNA-Seq transcript data can be compared with a closest annotated genome to obtain useful information [3,12]. This approach will also be helpful in obtaining sequence information of complex genomes, such as polyploid plants, for which genome sequencing is complex and time consuming, even with the latest sequencing methods [3].

Although RNA-Seq remains in the development stage, it has quickly established itself as an indispensable tool in biological research [11], and is expected to gain wider applications as developments ensue. At present, the cost of RNA-Seq is the major barrier to its widespread adoption [13]; however, it should drop as use of the method increases. The development of cost-effective and efficient library construction protocols or strategies to sequence RNA without needing to produce cDNA, as well as increasing the power to capture rare transcript species are likely to remain the targets for future research. In common with other NGS technologies, the reduction of cost of RNA-Seq would generate huge amounts of data, leading to challenges of archiving, retrieving, and processing the data. These challenges should also be addressed in order to make RNA-Seq a more efficient and applicable tool.

## References

1. Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci U S A 74: 5463-5467.
2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, et al. (2001) The sequence of the human genome. Science 291: 1304-1351.
3. Schuster SC (2008) Next-generation sequencing transforms today's biology. Nat Methods 5: 16-18.
4. McCutcheon JN, Giaccone G (2015) Next-Generation Sequencing: Targeting Targeted Therapies. Clin Cancer Res .
5. Metzker ML (2010) Sequencing technologies - the next generation. Nat Rev Genet 11: 31-46.
6. de Vree PJ, de Wit E, Yilmaz M, van de Heijning M, Klous P, et al. (2014) Targeted sequencing by proximity ligation for comprehensive variant detection and local haplotyping. Nat Biotechnol 32: 1019-1025.
7. Chung CH, Guthrie VB, Masica DL, Tokheim C, Kang H, et al. (2015) Genomic alterations in head and neck squamous cell carcinoma determined by cancer gene-targeted sequencing. Ann Oncol 26: 1216-1223.
8. Jour G, Benayed R, Reilly C, Ladanyi M, Arcila M (2015) RNA based next generation sequencing (NGS) assay to detect gene fusions in translocation associated sarcomas. In: Laboratory Investigation. Nature.
9. Bombarely A, Coate JE, Doyle JJ (2014) Mining transcriptomic data to study the origins and evolution of a plant allopolyploid complex. PeerJ 2: e391.
10. Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nat Rev Genet 10: 57-63.
11. Creecy JP, Conway T (2015) Quantitative bacterial transcriptomics with RNA-seq. Curr Opin Microbiol 23: 133-140.

12. Ohtsu K, Smith MB, Emrich SJ, Borsuk LA, Zhou R, et al. (2007) Global gene expression analysis of the shoot apical meristem of maize (Zea mays L.). Plant J 52: 391-404.

13. Hou Z, Jiang P, Swanson SA, Elwell AL, Nguyen BK, et al. (2015) A cost-effective RNA sequencing protocol for large-scale gene expression studies. Sci Rep 5: 9570.