The Necessity of Using Strand-Specific cDNA for Achieving Accurate Transcriptome Analysis Result

Dear Editor,

The importance of transcriptome analysis results due to their critical roles in genome annotation and subsequently progression in prevention, prognosis, diagnosis, and treatment of disease, led us on sharing our concern about the accuracy level of these existent data and meta-analysis on them. Transcriptome analysis, especially RNA sequencing (RNA-seq), has increasingly provided new insights for understanding gene structure, expression, and regulation and has been developed in some methodology.[1,2] There are regions in the genome (nearly 10% of the human genome loci) that produce both coding and noncoding isoforms that are called bifunctional RNAs and were confirmed and annotated by NCBI. The balance between noncoding and coding RNAs levels are affected by both physiologic process like development stages and differentiation and also environmental factors such as drug, physical or chemical agent or viral, bacterial or fungal pathogens.[3] Moreover, information in antisense transcripts plays important role in transcriptome profiling according to their critical roles in biological functions, as well as, by affecting accurately quantifying sense gene expression, particularly for genes with the overlapping loci and opposite direction of transcription. Therefore, discrimination between coding and noncoding isoforms is crucial to prevent misinterpretation of RNA-Seq data.

Comparative evaluations have shown that according to used protocol for cDNA library preparation, sometimes different results are obtained from the same samples. Strand information is vital to analysis transcriptome and if cDNA libraries are prepared by nonstrand-specific method, all overlapped sense and antisense strands including coding and noncoding RNA will get quantified as a "sense" signal, then distinguish between antisense and novel RNA species (for example LncRNA) remains problematic and can lead to inaccurate evaluation of sense RNA expression level.[2,4,5]

To deal with this problem, some approaches are developed including attaching various adapters to the 5′ and 3′ ends of the RNA molecules, bisulfite treatment, incorporating Deoxyuridine triphosphate (dUTP) into the second strand of cDNA and adding adapters to fragmented RNA. Despite the advantages and disadvantages of these methods, commercial directional RNA-Seq library preparation protocols utilized them.[2,6] However, different strand-specific protocols have represented different results in genome annotation and expression profiling as well.[7] and both strand-specific and nonstrand methods are used frequently. Altogether, these facts can hazardously impact on the results of interpretation of meta-transcriptomic data which performed on studies with different preparation methods.[8]

According to our literature search, unfortunately, most meta-analysis studies on high throughput transcriptomics did not notice to this bias, so their results can be controversial. Therefore, there is an urgency to standardize data from different studies and notice the type of used library preparation protocol (strand-specific or nonstrand-specific RNA sequencing) before comparison to obtain accurate results. Although quantitative real-time reverse transcription PCR can be used to validate transcript abundance, the bias introduced at this stage has not been sufficiently addressed.[8]

Indeed, it is thought that priming of reverse transcription using gene-specific primer for a specific transcript leads to the production of cDNA from only that required RNA template and guarantees the specificity of the signal in subsequent PCR amplification. However, studies have shown that cDNA can be synthesis without adding an exogenous primer due to the self-priming properties of the 3′ end of RNA template[9] and lead to the production of nonspecific cDNAs via a phenomenon term as priming background.[10] Primer-independent cDNA synthesis impairs the quantitative analysis of bidirectional transcripts when both strands are co-expressed in eukaryotes and cells infected with RNA viruses as well.[9] On the other hand, background priming not only interferes with the accurate evaluation of specific transcripts, but also prevents correct measurement of the effectiveness of knockdown during RNA interference (RNAi) investigations to reduce the function of particular transcripts.[11]

Many methods were developed to overcome this problem, including RNase-H-mediated digestion of nontarget RNA strand, adding DMSO to RT reaction, using the thermostable reverse transcriptase, performing RT at a higher extension temperature, utilizing reverse transcriptase with RNase-H activity, the insertion of extra tags to RT primers and then detection of tagged cDNAs, altering the primer-specific cDNA's sequence, using Exonuclease-I to break down unused RT primers, purification of cDNA produced with biotinylated primers and denaturing cDNA-cDNA duplexes during biotinylated cDNA enrichment. However, each of these methods alone has not been very successful to detect the strand-specific cDNAs, and the few combinations of these methods have shown success with varying degrees.[11] Altogether, this evidence indicates the urgency of determining standard pipelines for transcriptome analysis in order to increase accuracy of expression profiling to ensure reproducibility of results across libraries and sequencing technologies.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

Reyhaneh Dehghanzad[1], Anis Khalafiyan[2], Hossein Khanahmad[2]

[1]Department of Medical Genetics, Faculty of Medical Science, Tehran University of Medical Science, Tehran, Iran, [2]Department of Genetics and Molecular Biology, Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

Address for correspondence: Dr. Hossein Khanahmad, Department of Genetics and Molecular Biology, Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran.
E-mail: hossein_khanahmad@yahoo.com

Submitted: 02-Apr-2022; **Revised:** 04-Aug-2022;
Accepted: 06-Aug-2022; **Published:** 27-Apr-2023

References

1. Wang B, Kumar V, Olson A, Ware D. Reviving the transcriptome studies: An insight into the emergence of single-molecule transcriptome sequencing. Front Genet 2019;10:384.
2. Hrdlickova R, Toloue M, Tian B. RNA-Seq methods for transcriptome analysis. Wiley Interdiscip Rev RNA 2017;8:10.
3. Hubé F, Francastel C. Coding and non-coding RNAs, the frontier has never been so blurred. Front Genet 2018;9:140.
4. Sarantopoulou D, Tang SY, Ricciotti E, Lahens NF, Lekkas D, Schug J, *et al*. Comparative evaluation of RNA-Seq library preparation methods for strand-specificity and low input. Sci Rep 2019;9:13477.
5. Zhao S, Zhang Y, Gordon W, Quan J, Xi H, Du S, *et al*. Comparison of stranded and non-stranded RNA-seq transcriptome profiling and investigation of gene overlap. BMC Genomics 2015;16:675.
6. Mills JD, Kawahara Y, Janitz M. Strand-Specific RNA-Seq Provides Greater Resolution of Transcriptome Profiling. Curr Genomics 2013;14:173-81.
7. Levin JZ, Yassour M, Adiconis X, Nusbaum C, Thompson DA, Friedman N, *et al*. Comprehensive comparative analysis of strand-specific RNA sequencing methods. Nat Methods 2010;7:709-15.
8. Alberti A, Belser C, Engelen S, Bertrand L, Orvain C, Brinas L, *et al*. Comparison of library preparation methods reveals their impact on interpretation of metatranscriptomic data. BMC Genomics 2014;15:912.
9. Tuiskunen A, Leparc-Goffart I, Boubis L, Monteil V, Klingström J, Tolou HJ, *et al*. Self-priming of reverse transcriptase impairs strand-specific detection of dengue virus RNA. J Gen Virol 2010;91:1019-27.
10. Frech B, Peterhans E. RT-PCR: 'Background priming' during reverse transcription. Nucleic Acids Res 1994;22:4342-3.
11. Uddin F, Srivastava M. Strand-specific detection of overlapping transcripts via purification involving denaturation of biotinylated cDNA. BioTechniques 2020;69:141-7.

Access this article online

Quick Response Code:

Website: www.advbiores.net

DOI: 10.4103/abr.abr_102_22

How to cite this article: Dehghanzad R, Khalafiyan A, Khanahmad H. The necessity of using strand-specific cDNA for achieving accurate transcriptome analysis result. Adv Biomed Res 2023;12:108.