Transcriptomics: Open Access

Cheng and Lim, Transcriptomics 2013, 2:1
http://dx.doi.org/10.4172/2329-8936.1000e106

Editorial Open Access

Organelle Transcriptomes in Plants

Shifeng Cheng and Boon L Lim*

School of Biological Sciences, University of Hong Kong, Hong Kong

Keywords: Plant organelles; RNA-seq; Posttranscriptional process; RNA editing

Researches on plant organelle genomes are intriguing due to their distinctive characteristics: both plastids and mitochondria are of bacterial endosymbiont origins with deconstructed genomes, presenting many enigmatic features gained/shared from both prokaryotes and eukaryotes in the evolution; chloroplasts and mitochondria are the two key power houses of plant cells and many components of the energy generating systems (e.g. photosystems in chloroplasts and respiratory chain in mitochondria) are encoded by both nuclear and organelle genomes; the cross-talk between nuclear and organelle gene transcription and translation is coordinately regulated with high specificity. One example is the exclusively land plant-expanded nucleus-encoded RNA-binding PPR (pentatricopeptide repeat) proteins which participate in specific RNA processing mechanisms in both organelles. All of these characteristics make the researches on plant organelles fascinating. Next generation sequencing technologies (NGS) has completed the nucleus and organelle genome sequences of many plant species, which were designated as the uncovering of "the book of life" of an organism. However, it is the transcription and dynamic coordination between nuclear and organelle genes that specify a plant cell, here we can refer to the elucidation of plant transcriptome as revealing "the legend of life".

Numerous studies on the global transcriptome of the model plant Arabidopsis adopted microarray technology in the past decades. However these studies either employed Affymetrix ATH1 microarray or homemade microarray with limited number of probes. The Affymetrix ATH1 microarray only contains 22,500 probe sets for 24,000 genes and could not represent all the genes in the Arabidopsis genome (>30,000 genes in TAIR10), regardless of the various RNA spliced variants. Besides, no genes from the chloroplast and mitochondrial genomes are represented; the missing of transcription data of organelle genomes is a major deficiency of these studies. While the microarray chip from another supplier (NimbleGen) contains >30,000 probe sets that represent the whole nuclear genome and partial organellar genomes, the expression data on transcripts encoded by organellar genomes should be interpreted in caution due to the complex RNA sequences produced by posttranscriptional processes which still could not be differentiated by these microarray studies.

Although the genomes of these two organelles are small and contain relative few genes, the RNA transcribed from organellar genomes undergo extensive posttranscriptional processing, including 5'-end and 3'-end processing, multicistronic transcription, intercistronic processing, RNA splicing and RNA editing [1]. RNA editing is a unique feature of land plant plastids and mitochondria, which involves the conversion of specific cytosine (C) to uracil (U) nucleotides after mRNA transcription in seed plants. The conversion is not due to a substitution of nucleotide but rather is generated by a deamination or transmination reaction [2]. 41 and 441 editing sites in the Arabidopsis chloroplast and mitochondrial genomes have been identified by the traditional Sanger sequencing technology [3,4]. Some of these editing sites result in amino acid substitutions (e.g. Ser→Leu, Thr→ Met, etc) and the degree of RNA editing of some sites were shown to be tissue- and/or development-dependent [5], thus the qualification and quantification of RNA editing efficiency responded to different internal/external environments is an open question for scientists. All of the above posttranscriptional processes produce complex RNA sequences that could not be differentiated by microarray. RNA sequencing (RNA-seq) is thus a more powerful technique for studying the transcriptome of these two organelles.

To sequence organellartranscripts by RNA-seq, the experimental method for preparing RNA should be carefully considered. The reason is, organellar transcripts do not generally contain polyA tail, as do the nuclear transcripts, and posttranscriptional polyadenylationof organellar transcripts accelerates their degradation [6]. Therefore, the general approach of isolating poly(A)+ mRNA by oligo (dT) before cDNA synthesis will lead to biases. Removal of ribosomal RNA by rRNA removal kit is the preferred methodfor the studies on organellar transcriptome.

We have been studying fast-growing transgenic Arabidopsis which contained significantly higher leaf sucrose [7,8] and ATP [9] than the wild-type (WT). Microarray studies showed that transcriptional regulation does play important roles in starch, sucrose, N, K and Fe uptake, amino acids and secondary metabolites metabolisms [9]. By adopting the RNA-Seq approach described above, total leaf transcriptome profiles between the WT and the fast-growing transgenic plant were compared at three time points (t=0, 1, and 8 hr after illumination). The transcripts of 29,435 genes were detected in the six datasets and eleven to thirteen thousands alternative splicing events were identified in each dataset, all of which can further improve the Arabidopsis genome annotation. Moreover, transcripts encoded by the chloroplast and mitochondrial genomes were quantified and the degree of RNA edit were compared between the plants. Our data show that RNA-seq is a powerful tool for studying plant organellar transcriptome. More findings on the regulation of plant organellar transcriptome under various conditions (e.g. light intensity, wavelength, and stresses) will provide more information on plant physiology.

Nonetheless, the "standard" RNA-seq methods described in many researches are often not adequate. The major reason is that high-throughput RNA-seq derived from current NGS technology require extensive amplification process for short read sequences, which actually are prone to sequencing errors (about one error in one thousand nucleotides (0.1%) in current popular high-throughput sequencing platforms [10]. To abate such errors, besides to increasing sequencing depths followed by bioinformatics filtering, more biological replicates or cross-platform replicates are required. In the calls of RNA editing sites by deep RNA-seq, the center of controversy is exactly the sequencing and mapping errors and genetic variants. Li

***Corresponding author:** Boon L Lim, School of Biological Sciences, University of Hong Kong, Hong Kong, Tel: 2299 0826; E-mail: bllim@hkucc.hku.hk

Received December 30, 2013; **Accepted** January 02, 2014; **Published** January 06, 2014

Citation: Cheng S, Lim BL (2014) Organelle Transcriptomes in Plants. Transcriptomics 2: e106. doi:10.4172/2329-8936.1000e106

et al. [11] reported ~ 10,000 RNA-DNA sequence differences (RNA editing events) in human cell, whereas most of those sites are argued to be false discovery [12]. Peng et al. [13] also described the deep RNA-seq of a Han Chinese individual and identified 22,688 RNA editing events across the genome. In our Arabidopsis RNA-seq study, after a very stringent bioinformatics filtering process, we identified in total 207 RNA editing sites in mitochondria, 163 of which are consistent with previously reported sites generated by Sanger sequencing [4], which indicates there is at least ~20% of these possible RNA editing sites revealed by RNA-seq need further validation.

To solve the above problems, many specialized RNA-seq technologies have been developed, especially for studying the complex post-transcription processes in organelle-derived transcripts and non-coding RNAs. Significantly, strand-specific RNA-seq approach is optimally applicable for RNA editing studies, because this method can preserve DNA strand information that was lost in the cDNA preparation process in most of the other RNA-seq methods [14]. More recently a modified strand-specific RNA-seq method with high mapping and quantifying rates: strand- and transcript-specific RNA-seq method (STS-PCRseq)was reported [15]. Another challenge for plant organelle transcriptome researches is the ubiquitous of homologous genes across organelles and nucleus: NUMTs (nuclear mitochondrial transcripts) and NUPTs (nuclear plastid transcripts) [16]. 48 and 17 out of 122 and 88 protein coding genes in Arabidopsis mitochondrial and chloroplast genomes have homologs in the nucleus. We should always keep in mind that when mapping total transcriptstome these homologous genes, it is often difficult to discriminate organelle-derived transcripts from nucleus-derived ones. Some researchers tried bulked segregated RNA-seq method [17], which is able to an extent to identify and quantify homologous genes with extensive polymorphism. While the bulked segregated RNA-seq experimental approach can improve this problem, much effort should be paid in computational algorithms for short reads mapping and quantifying (such as for the SOAP and also Cufflinks software, they usually abandon multiple-mapping reads and only keep unique mapping reads for expressed genes identification and quantitation). Besides, some other adapted RNA-seq methods have been designed, likedouble-stranded RNA-seq, which can elucidate the secondary structures of RNA [18], and differential RNA-seq, which can uncover the TSSs and many transcripts processed sites [12]; the later method focus on non-coding RNA, and can describe an entire RNA molecule from 5'- to 3'- end.

In general, short-read RNA sequencing has been widely employed in the description of gene expression. However, it is still limited in the aspects of specificity and sensitivity especially for RNA editing qualification and quantification, which can only be inferred from a patchwork of short fragments. Similarly, RNA-seq is helpful but is still limited for the mapping of homologous genes; moreover, most RNA-seq techniques cannot identify full-length transcript isoforms using short reads. In the future, besides reducing the sources of errors by increasing sequencing read depth followed by optimized bioinformatics filtering steps, or by adding more technical, biological, or cross-platform replicates for one RNA-seq experiment, two strategies can improve the problems described above in plant organelles transcriptome researches: 1, plant organelles transriptome studies, that is, extract transcripts from isolated mitochondria or chloroplast for RNA-seq; 2, the newly single-molecule long-read sequencing technology from Pacific Biosciences, of which the read length can reach up to 1.5 kbp without any fragmentation and amplification [19], there by resolving many potential problems met by current multiple RNA-seq methods.

References

1. Stern DB, Goldschmidt-Clermont M, Hanson MR (2010) Chloroplast RNA metabolism. Annu Rev Plant Biol 61: 125-155.
2. Shikanai T (2012) Why Do Plants Edit RNA in Plant Organelles?. In: Organelle genetics: Evolution of Organelle Genomes and Gene Expression, (ed) Bullerwell CE, Editor Springer-Verlag Berlin Heidelberg 381-397.
3. Tillich M, Funk HT, Schmitz-Linneweber C, Poltnigg P, Sabater B, et al. (2005) Editing of plastid RNA in Arabidopsis thaliana ecotypes. Plant J 43: 708-715.
4. Giege P, Brennicke A (1999) RNA editing in Arabidopsis mitochondria effects 441 C to U changes in ORFs. Proc Natl Acad Sci USA 96: 15324-15329.
5. Hirose T, Sugiura M (1997) Both RNA editing and RNA cleavage are required for translation of tobacco chloroplast ndhD mRNA: a possible regulatory mechanism for the expression of a chloroplast operon consisting of functionally unrelated genes. EMBO J 16: 6804-6811.
6. Kuhn J, Tengler U, Binder S (2001) Transcript lifetime is balanced between stabilizing stem-loop structures and degradation-promoting polyadenylation in plant mitochondria. Mol Cell Biol 21: 731-742.
7. Sun F, Suen PK, Zhang Y, Liang C, Carrie C, et al. (2012) A dual-targeted purple acid phosphatase in Arabidopsis thaliana moderates carbon metabolism and its overexpression leads to faster plant growth and higher seed yield. New Phytol 194 : 206-219.
8. Sun F, Carrie C, Law S, Murcha M W, Zhang R, et al. (2012) AtPAP2 is a tail-anchored protein in the outer membrane of chloroplasts and mitochondria. Plant Signal Behav 7: 927-932.
9. Sun F, Liang C, Whelan J, Yang J, Zhang P, et al. (2013) Global transcriptome analysis of AtPAP2 - overexpressing Arabidopsis thaliana with elevated ATP. BMC Genomics 14: 752.
10. Ratan A, Miller W, Guillory J, Stinson J, Seshagiri S, et al. (2013) Comparison of sequencing platforms for single nucleotide variant calls in a human sample. PLoS One 8: e55089.
11. Li M, Wang IX, Li Y, Bruzel A, Richards AL, et al. (2011) Widespread RNA and DNA sequence differences in the human transcriptome. Science 333: 53-58.
12. Zhelyazkova P, Sharma C M, Forstner K U, Liere K, Vogel J, et al. (2012) The primary transcriptome of barley chloroplasts: numerous noncoding RNAs and the dominating role of the plastid-encoded RNA polymerase. Plant Cell 24: 123-136.
13. Peng Z, Cheng Y, Tan BC, Kang L, Tian Z, et al. (2012) Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30: 253-260.
14. Ruwe H, Castandet B, Schmitz-Linneweber C, Stern DB (2013) Arabidopsis chloroplast quantitative editotype. FEBS Lett 587: 1429-1433.
15. Bentolila S, Oh J, Hanson MR, Bukowski R (2013) Comprehensive high-resolution analysis of the role of an Arabidopsis gene family in RNA editing. PLoS Genet 9: e1003584.
16. Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: the idiosyncratic genetics of endosymbiosis. Annu Rev Plant Biol 60: 115-138.
17. Liu S, Yeh CT, Tang HM, Nettleton D, Schnable PS (2012) Gene mapping via bulked segregant RNA-Seq (BSR-Seq). PLoS One 7: e36406.
18. Zheng Q, Ryvkin P, Li F, Dragomir I, Valladares O, et al. (2010) Genome-wide double-stranded RNA sequencing reveals the functional significance of base-paired RNAs in Arabidopsis. PLoS Genet 6: e1001141.
19. Sharon D, Tilgner H, Grubert F, Snyder M (2013) A single-molecule long-read survey of the human transcriptome. Nat Biotechnol 31: 1009-1014.

Citation: Cheng S, Lim BL (2014) Organelle Transcriptomes in Plants. Transcriptomics 2: e106. doi:10.4172/2329-8936.1000e106