



Category: Plant Genomics

De novo assembly and analysis of *Solanum trilobatum* L. leaf transcriptome using next generation sequencing technology

L. Adil and Purushothaman Natarajan*

Department of Genetic Engineering, SRM University, Kattankulathur 603203, INDIA

*Corresponding author: purushothaman.n@ktr.srmuniv.ac.in

Abstract

RNA Sequencing based *de novo* assembly is a well-developed approach in understanding transcriptomes of non-model plants with limited genomic information. RNA-Seq is cost effective tool, offers much data with better coverage and sufficient sequence depth for *de novo* assembly of transcriptomes. In past few years, there has been an increase in utilising RNA-Seq for discovery and identification of functional genes involved in the biosynthesis of active compounds in non-model plants. In this study, we analysed the transcriptome of *Solanum trilobatum* L. leaf using high throughput next generation sequencing. *S. trilobatum* is one of the important medicinal plants belonging to family Solanaceae and commonly available in South India. The studies conducted so far, to understand its therapeutic potential, have yielded positive results. Its extract is used to treat conditions like chronic bronchitis and tuberculosis. It is also reported to have anti-oxidative, hepatoprotective, anti-inflammatory, anti-microbial, anti-tumour activities. The total RNA from *S. trilobatum* leaf was isolated and sequenced using Illumina Hiseq 2500 platform with paired end chemistry. In total, 136,220,612 high quality sequence reads were obtained. The raw reads were pre-processed and assembled into 144,580 assembled transcripts using Trinity- a *de novo* assembler and clustering of transcripts was done using CD-HIT resulting 128,934 unigenes. The unigenes were extensively evaluated and annotated with various databases to identify pathways and genes responsible for biosynthesis of medicinal compounds. Based on similarity search with known proteins 60,097 (46.61% of all unigenes), 35,141 (27.25%), 30,427 (23.60%) and 61,986 (48.07%) had homologs in nr, Pfam, GO and UniProt databases respectively. The comparison against the KEGG database mapped 14,490 (11.23%) unigenes to 138 pathways, where flavonoid biosynthesis pathway was identified to be the highly represented. The expression levels of the transcripts were quantified using RSEM and Reverse Transcription PCR (RT-PCR) of few genes were performed to validate the transcriptome assembly. The SSRs and transcription factors, which could help for the molecular breeding, were also identified. This is the first report of complete transcriptome analysis in *S. trilobatum*. The genomic resources generated will serve as foundation to understand molecular basis of medicinal properties of *S. trilobatum* in further studies.

Citation: Adil, L. and Natarajan, P. *De novo* assembly and analysis of *Solanum trilobatum* L. leaf transcriptome using next generation sequencing technology [Abstract]. In: Abstracts of the NGBT conference; Oct 02-04, 2017; Bhubaneswar, Odisha, India: Can J biotech, Volume 1, Special Issue, Page 186. <https://doi.org/10.24870/cjb.2017-a172>