

## عنوان مقاله:

(.Different k-mer strategies and de novo assembly tools challenge for transcriptome analysis of Citrullus colocynthis (L

## محل انتشار:

چهارمین کنگره بین المللی و شانزدهمین کنگره ملی ژنتیک (سال: 1399)

تعداد صفحات اصل مقاله: 2

## نویسندگان:

Masoumeh Dorafshan - *Department of Genetics and Plant Breeding, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran*

Mehdi Soltani Howyzeh - *Department of Genetics and Plant Breeding, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran*

Vahid Shariati - *Department of Molecular Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran*

## خلاصه مقاله:

**Background and Aim:** De novo transcriptome assembly can obtain genetic information in a high throughput scale from non-model organisms like medicinal plants. But there is a rising challenge to select suitable transcriptome assemblers to perform a de novo assembly due to the fast developments of assembly software tools. In this research, to distinguish the best high quality de novo assembly strategy, we created a total of ۱۴ assemblies (with different k-mer lengths) from six different tissue samples of *C. colocynthis* as a non-model organism using different six recently published de novo assemblers including BinPacker, Bridger, Evidentialgene, Trinity, SOAPdenovo-Trans and SPAdes. **Methods:** The entire RNA was extracted using the Qiagen RNeasy kit (QIAGEN). Quality control for the extracted RNA was examined by using a QC Bioanalyzer (Agilent Technologies, Hørsholm, Denmark) and the RNA integrity number (RIN) of each sample was greater than ۸. The selection of Poly A, cDNA preparation, adapter ligation, the formation of clusters and sequencing was performed at the Beijing Genomes Institute (China), according to the manufacturer's recommendation, with the use of standard Illumina kits. The sequencing was done on an Illumina HiSeq ۲۵۰۰ platform with a paired-end and read length of ۱۵۰ nt. **Results:** The quality control processing resulted in ۱۲۰,۶۳۲,۵۰۴ high-quality filtered reads. Six quality metrics of de novo assembly were used to assess the de novo assemblies constructed by different strategies. These quality metrics were Total assembled bases, Transcript number, N۵۰ length, Average contig, Reads that can be Mapped Back to Transcripts (RMBT) and Benchmarking Universal Single-Copy Orthologs (BUSCO). None of the six assemblies accomplished the best in all metric categories tested. Trinity assembly strategy performance was within the top three of most metric categories, followed by the Bridger and BinPacker strategies, respectively. But Trinity assemblies were worst in BUSCO analysis due to the number of duplicated-copy BUSCOs in this research. In contrast to this, the Evidentialgene strategy was superior to others because of less fragmented and missing genes and more single-copy complete BUSCOs. **Conclusion:** Based on the evaluation metrics performed, the Evidentialgene strategy was able to construct the best assembly of the transcriptome in *C. colocynthis* medicinal plant.

## کلمات کلیدی:

.Medicinal Plant, RMBT, BUSCO, Trinity, Evidentialgene

