Int.J.Curr.Microbiol.App.Sci (2019) *8*(2): 3454-3458

International Journal of Current Microbiology and Applied Sciences
***ISSN: 2319-7706* Volume 8 Number 02 (2019)**
Journal homepage: http://www.ijcmas.com

Original Research Article https://doi.org/10.20546/ijcmas.2019.802.404

Analysis of GC-content in transcriptome sequence of *Coscinium fenestratum* (Gaertn.) Colebr Leaf

Ashalatha and S. M. Gopinath*

Department of Biotechnology, Acharya Institute of Technology, Bengaluru, Karnataka, India-560107

**Corresponding author*

Keywords

Guanine, Cytosine, Illumina paired end sequencing

Article Info

Accepted:
29 January 2019
Available Online:
10 February 2019

A B S T R A C T

Coscinium fenestratum (Gaertn.) Colebr, a important medicinal vine is considered as critically endangered or restricted to the humus rich soil. It has abundant use in ayurvedic, siddha, tibetan medicine system. The plant is already Red listed and at the verge of extinction. So a transcriptome study and the GC analysis of the plant are vital. The research provides information on its transcriptome and its stability which can be used for further studies. The leaf tissue of *C.fenestratum* was collected, sequenced using illumina paired end sequencing. The raw sequence data quality check parameters like the average base content and the GC content of the reads were analyzed. Maximum number of reads showed 43% of the average GC content in the sample showing slightly instability to adaptation.

Introduction

The ratio of four nitrogenous bases in nucleic acids may vary significantly in various genome components, its composition is conventionally expressed as the percentage of guanine (G) and cytosine (C) bases (GC content) in a given region or for the entire genome (genomic GC content). The study of GC serves as an important criterion in predictions of thermo tolerance, of the variety. However, less attention has been paid to analyze the GC content of plant genomes, for which the knowledge of detailed base composition and its meaning in the ecology and evolution of particular taxa is still poor.

The cause of variation in GC content is one of the central issues in evolutionary genomics. Some models link between GC content and temperature (Bernardi., 2000; Bernardi & Bernardi., 1986; Salinas *et al*., 1988). G: C pairs are more thermally stable than adenine (A) and thymine (T) pairs (Wada & Suyama 1986), G: C base pairs being are bonded by three hydrogen bonds and A: T base pairs by two. In turn, these interactions seem to be important in conferring stability to higher order structure for RNA transcripts (Smarda *et al*., 2012; Biro JC, 2008). A similar suggestion has been made for the evolution of

plant genomes (Salinas *et al*., 1988). So far, the highest GC contents of land plants have been found in grasses (Smarda *et al*., 2012; Smarda P et al., 2012; Salinas *et al*., 1988; Biswas SB, Sarkar AK 1970). In contrast to grasses, the lowest GC contents so far reported in plants Cyperaceae and Juncaceae (Lipneroval *et al*., 2013). By contrast, the GC content of Structural RNAs is higher at high temperatures.

Profound insight into the genomic architecture of model plants are rapidly accumulating, due to high-throughput next generation and third generation sequencing techniques (Flagel *et al*., 2012). However, the genomic constitution of the vast majority of nonmodel plants still remains unknown Genomic DNA base arrangement (GC content) is anticipated to essentially influence genome working and species adaptation to environment. The thermal theory demonstrates that genomic adjustments related with changing GC substance may have assumed a critical job in the development of the Earth's contemporary biota.

The reasons for the variation between genomes in their guanine (G) and cytosine (C) content is one of the focal issues in genomic studies. This GC and AT content variation is studied in *C.fenestratum* leaf transcript which is showing vulnerability in getting adapted to all climatic condition. And has been declared as critically endangered variety and show much adoption in Western Ghats or humid area (Ashalatha *et al*., 2019).

Materials and Methods

Plant Material

The leaf of *C.fenestratum* is collected and transferred to a RNA later solution to avoid RNA degradation.

RNA extraction

The RNA extraction of the leaf sample was carried out using RNAsol^{TM} Kit. The standard protocol provided was carried out for extraction.

RNA Purification

The extracted RNA was checked for 28S:18S RNA degradation by using an Agilent 2100 Bioanalyzer. The pooled RNA with an RIN (RNA integrity number) of 7.0 was used for further mRNA purification process. The obtained mRNA was further purified by oligo-dT beads using TruSeq kit.

Sequencing

The cDNA library was prepared and further the template was sequenced by a standardized protocol of Illumina paired end sequencing (Illumina Hi Seq 2500 platform, USA), with a read length of 101 * 2 by utilizing paired-end sequencing chemistry technique. The reads having ≥70% of the bases with a quality score ≥Q20 using NGS QC Toolkit [83] were chosen for assembling the transcriptome.

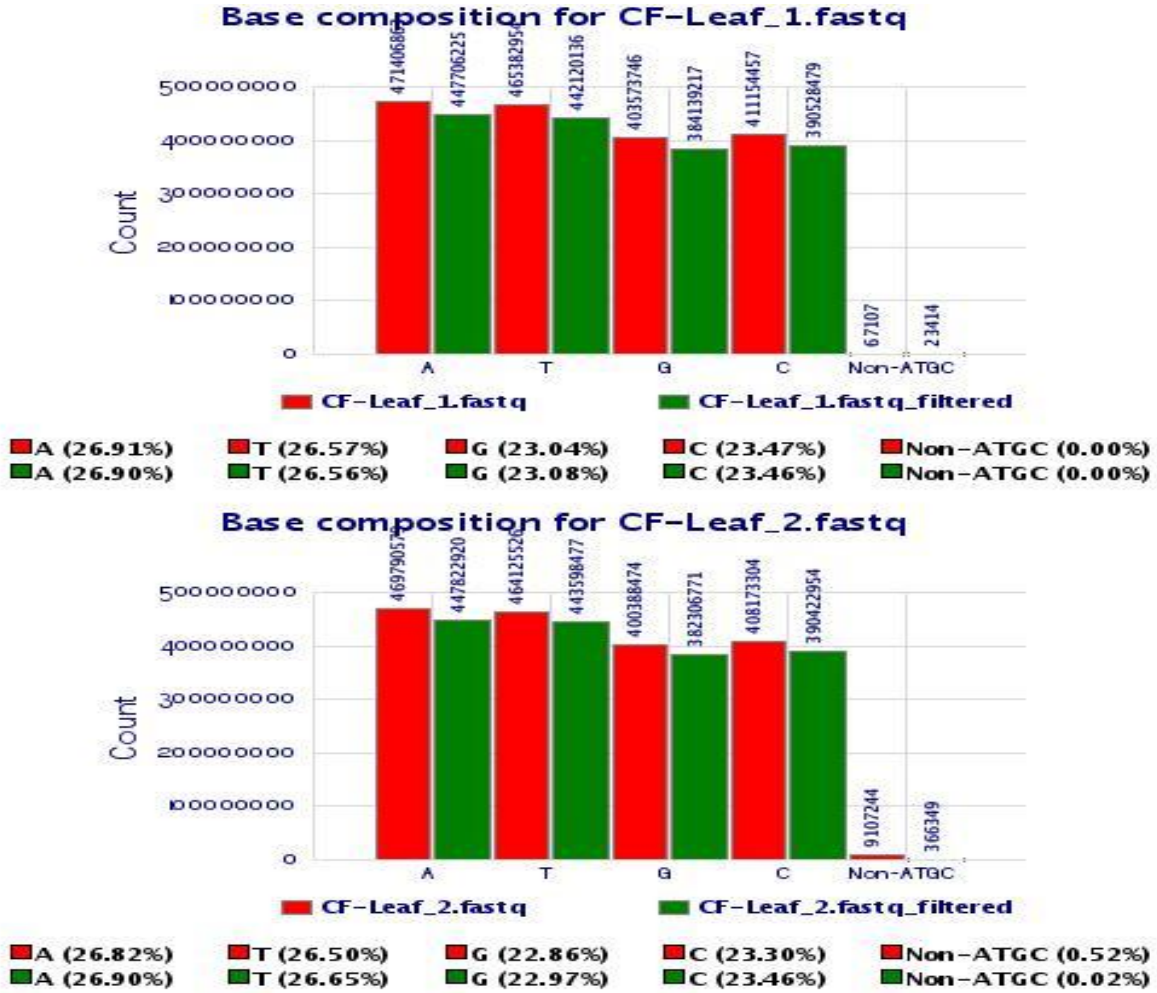
Results and Discussion

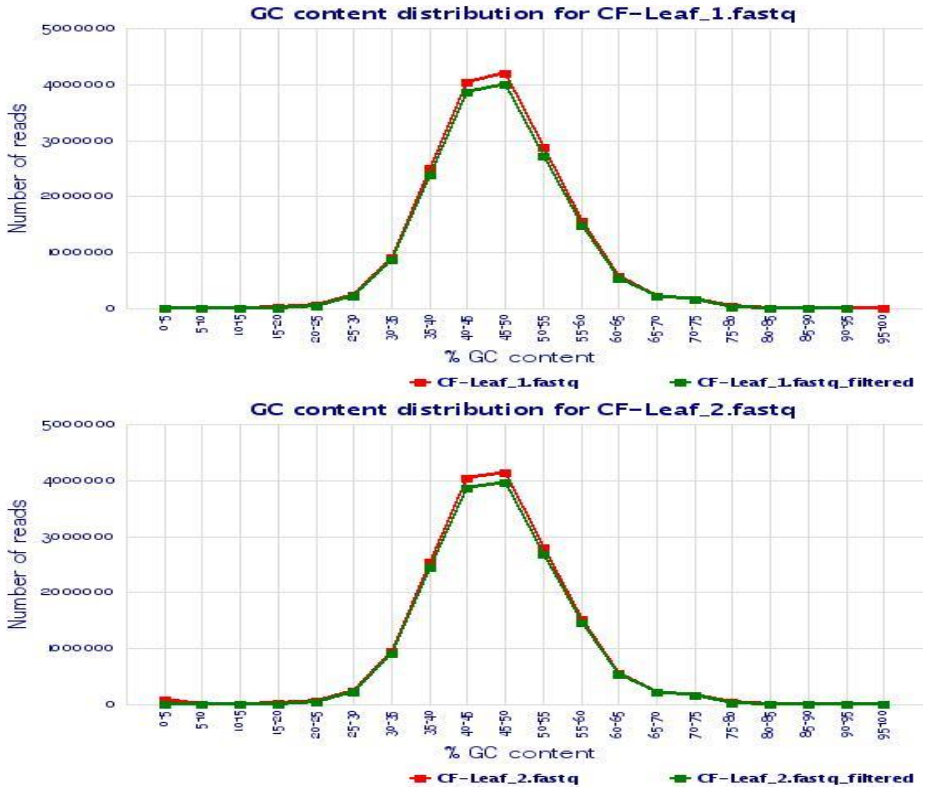
The present study was obtained an average of 17,342,427 total number of reads owing for 1,751,585,127 number of bases. The raw sequence data was deposited to the NCBI BioProject database (as SRA- Short Read Archive) with the accession number PRJNA415708. The other quality check parameters like the average base content and the GC content of the reads were analyzed. Maximum number of reads showed 43% of the average GC content in the sample. The reads in the samples follows the normal distribution of the GC content, which is similar to the theoretical GC distribution authenticating the quality of transcript

obtained. The data quality obtained is very good with 95.54% high quality reads the summary of GC content distribution (Figure: 1) of leaf is provided below.

The results showed 43% of the average GC content and 57% of AT content in the sample. This can be inferred as a sparse amount of thermal instability faced by the plants due to slightly high amount of A: T content. The adaptivity of the plant to all environmental condition is thus low (Franchi G. G., *et al* 2011).

Figure.1 GC content distribution of *Coscinium fenestratum* leaf





Acknowledgements

I am thankful to the Team of Bionivid Technologies, Bangalore for providing me the technical support and assistance to carry out my research work.

References

Ashalatha and S.M.Gopinath. (2019). "Phytochemical profiling of *Coscinium fenestratum (Gaertn.) Colebr*, by Liquid chromatography-Mass spectrometry". Int. J. Curr. Microbiol. App. Sci. 8(1):3194-3201.

Bernardi, G. (2000). Isochores and the evolutionary genomics of vertebrates. Gene 241, 3-17.

Bernardi, G. & Bernardi, G. (1986) Compositional constraints and genome evolution. J. Mol. Evol. 24, 1-11.

Biro, J.C. (2008). Correlation between nucleotide composition and folding energy of coding sequences with special attention to wobble bases. Theor Biol Med Model 5:14.

Biswas, S.B., Sarkar, A.K. (1970). Deoxyribonucleic acid base composition of some angiosperms and its taxonomic significance. Phytochemistry 9(12):2425–2430.

Franchi, G.G., Nepi, M., Dafni., A, Pacini, E (2002). Partially hydrated pollen: Taxonomic distribution, ecological and evolutionary significance. Plant Syst Evol 234(1-4):211–227.

Lipnerová, I., Bureš, P., Horová, L., Šmarda, P. (2013). Evolution of genome size in *Carex*(Cyperaceae) in relation to chromosome number and genomic base composition. Ann Bot (Lond)111(1):79–94.

Salinas, J., Matassi, G., Montero, L.M., Berna rdi, G (1988). Compositional compart-mentalization and compositional patterns in the nuclear genomes of plants. Nucleic Acids Res 16(10):4269–4285.

Wada, A. & Suyama, A. (1986). Local stability of DNA and RNA secondary

structure and its relation to biological functions. Prog. Biophys. Mol. Biol. 47, 113-157.

Šmarda, P., Bureš, P (2012) The variation of base composition in plant genomes. *Plant Genome Diversity*,eds Wendel F, Greilhuber J, D oležel J, Leitch IJ (Springer, Vienna), 1, 209–235.

Šmarda, P., Bureš, P., Šmerda, J., Horová, L. (2012). Measurements of genomic GC content in plant genomes with flow cytometry: A test for reliability. New Phytol 193(2):513–521.

Flagel,L.E., Blackman, B.K(2012).The first ten years of plant genome sequencing and prospects for the next decade. Plant Genome Diversity, eds Wendel JF, Greilhuber J, Dolezel, Leitch IJ (Springer, Vienna), 1,1-15

How to cite this article:

Ashalatha and Gopinath, S. M. 2019. Analysis of GC-content in transcriptome sequence of *Coscinium fenestratum* (Gaertn.) Colebr Leaf. *Int.J.Curr.Microbiol.App.Sci.* 8(02): 3454-3458. doi: https://doi.org/10.20546/ijcmas.2019.802.404