

## A Comparison Study of Chloroplast Genomes: Ficus Species Versus Its Close Relatives

Bihter Avsar<sup>1\*</sup> Danial Esmaceli Aliabadi<sup>2</sup>

<sup>1</sup>Sabancı University, Nanotechnology Research and Application Centre, Tuzla, 34956, Istanbul, Turkey

<sup>2</sup>Sabancı University, Faculty of Engineering and Natural Sciences, Tuzla, 34956, Istanbul, Turkey

\*Corresponding Author

E-mail: bihteravsar@sabanciuniv.edu

Received: September 18, 2018

Accepted: December 5, 2018

### Abstract

*Ficus* genus, mostly known as fig trees, contains about 850 species in the Moraceae family. This genus is vital due to its agronomical properties that the “fig” fruits are essential sources of flavonoids, sugars, vitamins A and C, acids and enzymes. *Ficus* genus is also an important part of the biodiversity of the rainforest ecosystem. Also, the previous studies showed that *Ficus racemosa* was found as one of the important medicinal plants for various diseases including diabetes, respiratory and urinary problems.

In our study, we decided to analyze a chloroplast genome from an important *Ficus* species: *Ficus racemosa*. We re-assembled the chloroplast genome of *Ficus racemosa* which results of 160,383 bp with a typical circular structure, and we found that the chloroplast genome contains total of 82 protein-coding genes, 56 tRNAs, and seven rRNA genes. Additionally, we also analyzed the chloroplast genomes of *Ficus carica*, *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba* for the further comparative studies. Repeat elements, Simple Sequence Repeats (SSRs), and phylogenetic tree analyses were conducted on the chloroplast genomes of all organisms and the results were compared. These findings presented in this study shows the level of chloroplast genome conservation, and the results will be discussed to provide clues for the understanding of the evolution of organellar genomes in tree species.

**Key words:** *Ficus*, *Ficus carica*, chloroplast, *Ficus racemosa*

### INTRODUCTION

According to the endosymbiotic theory, chloroplasts originated from photosynthetic organisms and retain their own unique DNA encoding multiple genes and is responsible for photosynthesis to convert light energy into chemical energy. In plants, chloroplast genomes have a conserved genome structure and the size is ranging from 120 kb to 160 kb [1]. Although, in both mitochondrial and nuclear genomes, the random and large size variations are more frequent, in chloroplast genomes, these variations are limited and are rarely detected. The limited size change in chloroplast genomes indicates the possibility that the chloroplast genome is maintained by natural selection [2].

*Ficus racemosa* is a species in the family of Moraceae family and it is well known as the cluster fig tree, Indian fig tree or goolar (gular) fig which is native to Australia, Southeast Asia, and the Indian subcontinent. *Ficus racemosa* (Moraceae) is an evergreen, moderate to large sized spreading, lactiferous, deciduous tree and it is an important element of lowland rainforest [3]. The fruits of the tree are edible and both fruits and the leaves are used in the treatment of diarrhea, haematuria, menorrhagia, and haemoptysis. Although *Ficus racemosa* has a high value for medicinal uses, genetic studies of these species and their relatives are limited. Therefore, we used genomic sequence information to generate the complete chloroplast sequence and we compared other *Ficus* species.

In one of the study, they reported a chloroplast genome of *Ficus racemosa* using different sequencing and assembly method [4]. In this study, we reported the complete chloroplast genome of the *Ficus racemosa* using Illumina reads with a different assembly method. We also analyzed the chloroplast genomes of *Ficus carica*, *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba* for the further comparisons. These results are important for understanding the complexity of chloroplast genomes and point out the dependence be-

tween chloroplast size, structure, the number of the genes and the ecological diversity and environmental complexity.

### MATERIALS and METHODS

#### Plant materials, chloroplast assemblies and genome annotations:

For *Ficus racemosa*, genomic DNA was fragmented and sequencing libraries were prepared using the TruSeq DNA Library Preparation kit (Illumina). Paired-end sequencing was carried out using a HiSeq4000 instrument (Illumina) with 76 nucleotide read lengths (SRX601830). For *Ficus carica*, genomic DNA was fragmented and the sequencing libraries were prepared using the TruSeq DNA Library Preparation kit (Illumina) [5]. Paired-end sequencing was carried out using a HiSeq4000 instrument (Illumina) (NCBI Reference Sequence: NC\_035237.1). For *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba*, complete chloroplast genomes were retrieved from NCBI. Their reference sequences are given as NC\_016736.1, NC\_007957.1, and KX266829.1, respectively.

Chloroplast assembly was performed by NovoPlasty 2.4 [6]. Once we generated an assembly, we used GeSeq online annotation tool [7] to find out the putative genes, gene boundaries, coding sequences, tRNAs and rRNAs. During this analysis, the close relatives were used as reference organisms. Chloroplast map was drawn by OGDRAW [8].

#### Simple sequence repeats (SSRs) and other repeat structures on chloroplast genomes

We detected SSRs for all chloroplast genomes by MISA Perl script [9] with a motif size of 1 to 6 nucleotides and minimum repeats size for each motif was set as 8, 4, 3, 3 and 3, respectively. Forward, reverse and palindromic repeat structure were identified and located on the chloroplast genome by REPuter [10] with default settings including hamming distance: 3, maximum computed repeats:50, minimal repeat size: 8. All the identified repeat structures were con-

firmed by various programs and the suspicious results were removed.

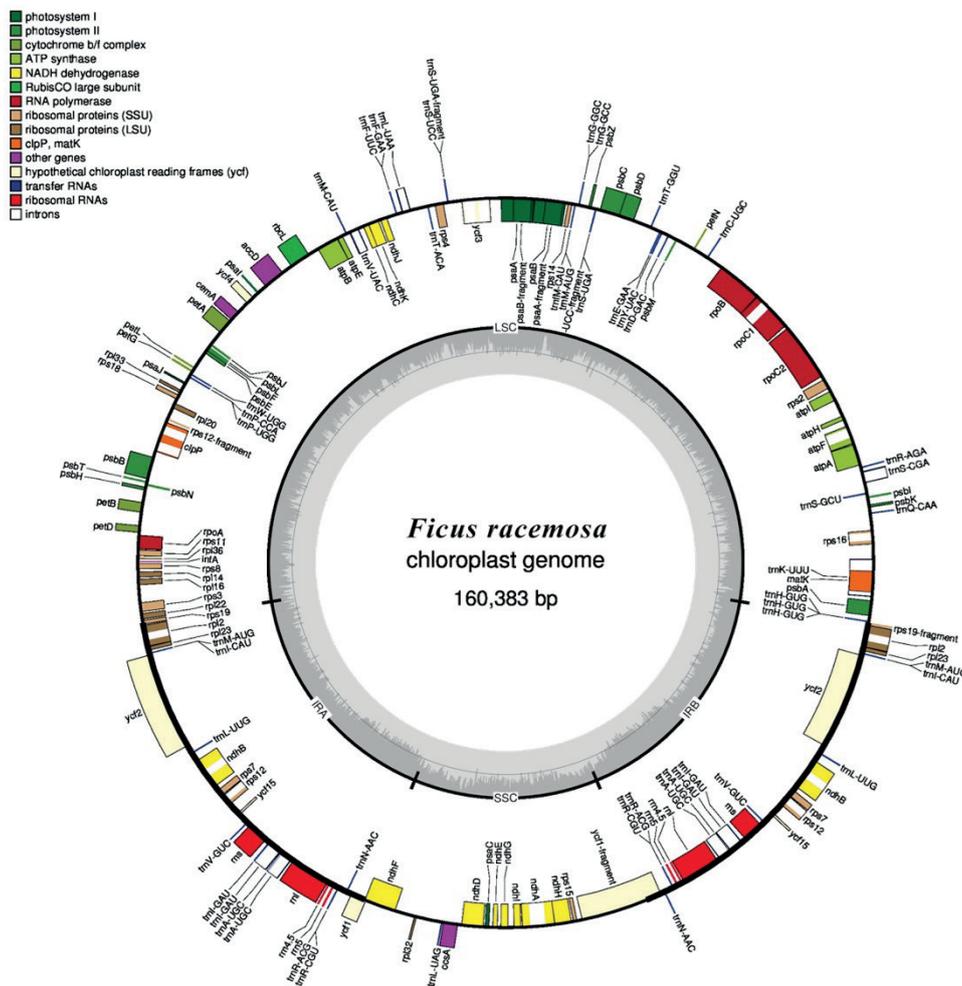
**Phylogenetic analysis**

We found 77 common chloroplast coding sequences from *Ficus racemosa*, *Ficus carica*, *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba*. These coding sequences were accD, atpA, atpB, atpE, atpF, atpH, atpI, ccsA, cemA, clpP, matK, ndhA, ndhB, ndhC, ndhD, ndhF, ndhG, ndhH, ndhI, ndhJ, ndhK, petA, petB, petD, petG, petL, petN, psaA, psaB, psaC, psaI, psaJ, psbA, psbB, psbC, psbD, psbE, psbF, psbH, psbI, psbJ, psbK, psbL, psbM, psbN, psbT, psbZ, rbcL, rpl14, rpl16, rpl2, rpl20, rpl22, rpl23, rpl32, rpl33, rpl36, rpoA, rpoB, rpoC1, rpoC2, rps11, rps12, rps14, rps15, rps16, rps18, rps19, rps2, rps3, rps4, rps7, rps8, ycf1, ycf2, ycf3, ycf4. All the genes were aligned using MUSCLE in MEGA6 software and were manually adjusted. We constructed the phylogenetic tree using Maximum-likelihood method in MEGA6 [11] and the branch point confidence was estimated based on 1000 bootstrap replications. The phylogenetic results were confirmed by other methods including Neighbour-joining and Maximum parsimony.

**RESULTS and DISCUSSIONS**

**Chloroplast assembly of *Ficus racemosa***

The size of the chloroplast genome of *Ficus racemosa* was 160,383 bp. There are 82 protein-coding genes, seven rRNA genes and 56 tRNA genes in the *Ficus racemosa* chloroplast genome. A total of 78 unique protein-coding genes, unique 21 tRNA genes, and seven unique rRNA genes were identified. ndhB, rpl23, rps7, and ycf2 protein-coding genes were found as two copies in the chloroplast genome. Inverted repeat regions, large single copy, and short single copy regions were also specified on the map (Figure 1). *Ficus racemosa* chloroplast genome consists of base A (31.68 %), base T (32.43%), base G (17.65%) and base C (18.24%) nucleotides. GC content is 35.89%, however AT content is 64.11%.



**Figure 1.** A representative circular image of *Ficus racemosa* chloroplast genome. Genes belong to different functional groups are shown in different color codes. GC content is represented on the inner gray circle. Gene boundaries, two inverted repeat (IR) regions (IRA, IRB), Short Single Copy(SSC) and Large Single Copy(LSC) parts were also shown in the image.

**Repeat structures in chloroplast genomes of *Ficus racemosa*, *Ficus carica*, *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba***

SSRs were found in all related chloroplast genomes (Figure 2). SSR structures of *Ficus carica*-*Ficus racemosa*

and *Ricinus communis*-*Vitis vinifera* are similar and this may be caused from the similarity in their chloroplast genomes. *Ricinus communis*, which has the largest chloroplast genome among the organisms in this study, has also the highest number of SSRs (330 identified SSRs) in the chloroplast genome.

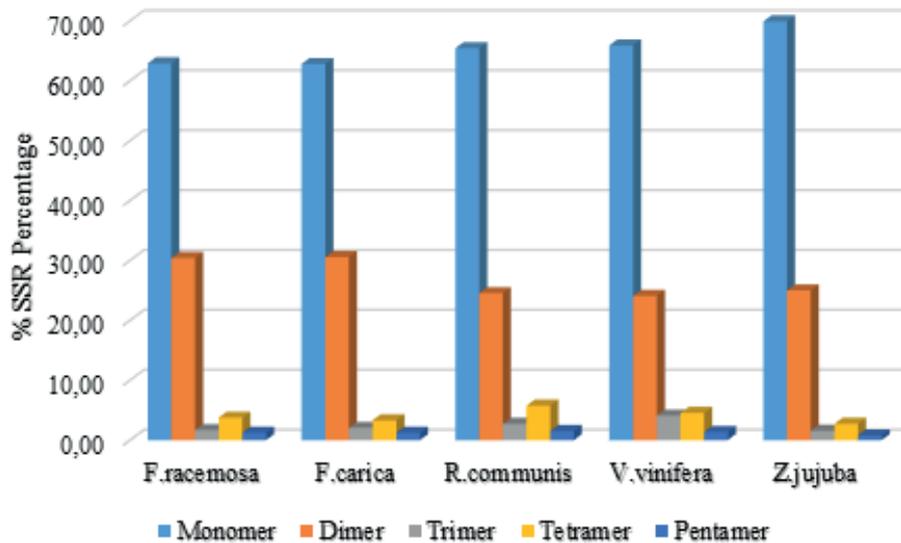


Figure 2. % SSR percentage of all chloroplast genomes.

REPuter software offers four possible repeats in the genome: forward (direct) match, reverse match, complement match, and palindromic match. In *Ficus racemosa* chloroplast genome, we could not find complement match type of

repeat structure, however other types of repeats were found in percentage (Figure 3). Palindromic repeats were found predominantly in *Ficus racemosa* chloroplast genome.

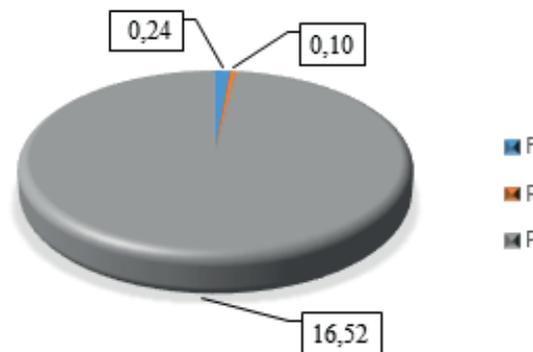
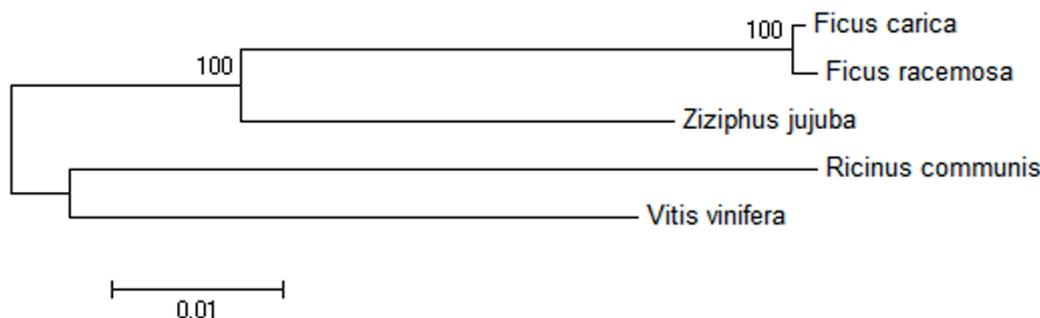


Figure 3. REPuter results of *Ficus racemosa* chloroplast genome. F: Forward repeats, R: Reverse repeats, P: Palindromic repeats.

**Phylogenetic tree analysis**

We conducted phylogenetic tree analysis in *Ficus racemosa*, *Ficus carica*, *Ricinus communis*, *Vitis vinifera* and *Ziziphus jujuba* using MEGA6 with Maximum-likelihood method based on 1000 bootstrap estimations. As illustrated in Figure 4, *Ficus racemosa* is evolutionarily closely related to *Ficus carica*.



**Figure 4.** The Maximum-likelihood phylogenetic tree was constructed based on 77 common protein-coding genes from five plant species (*Ficus carica*, *Ficus racemosa*, *Ziziphus jujuba*, *Ricinus communis* and *Vitis vinifera*).

## CONCLUSIONS

In this study, we used a newly developed organelle assembly software, Novoplasty 2.4, using Illumina paired-reads and we successfully were able to obtain a chloroplast genome map of *Ficus racemosa*. We also identified close relatives of this organism using Blast approach. Then, we did a comparison between the plant species based on their repeat structures and as expected, *Ficus* species showed high similarity to each other, whereas, among other species, *Ricinus communis* and *Vitis vinifera* were similar. Phylogenetic tree results gave consistent results, and evolutionarily, *Ficus* species were more closely related than the other species.

In future studies, mitochondrial and chloroplast genome comparison of *Ficus* species contributes to the plant evolution. Besides, having organellar genome information can be helpful for abiotic stress signaling in plants [12], as well as identification of microRNA targets on the chloroplast genome might provide us a new point of view of development, growth and stress mechanisms [13, 14, 15]. Furthermore, complete organellar genomes would provide clues for better understanding intra and inter genomic DNA transfers and recombination events. In addition to these, by using organelle genomes, marker development could be useful for breeding programs to identify agronomically important properties in plants.

### Acknowledgments

We acknowledge to Daniel Lee Calvey for his professional English writing of the manuscript and providing valuable comments for its enrichment.

## REFERENCES

- [1] Palmer, J. D. (1985). Comparative organization of chloroplast genomes. *Annual review of genetics*, 19(1), 325-354.
- [2] Xiao-Ming, Z., Junrui, W., Li, F., Sha, L., Hongbo, P., Lan, Q., Jing, L., Yan, S., Weihua, Q., Lifang, Z., Yunlian, C. (2017). Inferring the evolutionary mechanism of the chloroplast genome size by comparing whole-chloroplast genome sequences in seed plants. *Scientific reports*, 7(1), 1555.
- [3] Varier FS (1995): Indian Medicinal Plants, Vol. 3. Hyderabad, India, Orient Longman, pp. 34-37.
- [4] Mao, Q., Bi, G. (2016). Complete chloroplast genome of *Ficus racemosa* (Moraceae). *Mitochondrial DNA Part A*, 27(6), 4655-4656.
- [5] Rabah, S. O., Lee, C., Hajrah, N. H., Makki, R. M., Alharby, H. F., Alhebshi, A. M., Sabir, A.M., Jansen, J.S., Ruhlman, T. A. (2017). Plastome sequencing of ten nonmod-

el crop species uncovers a large insertion of mitochondrial DNA in cashew. *The plant genome*.

[6] Dierckxsens, N., Mardulyn, P., Smits, G. (2016). NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic acids research*, 45(4), e18-e18.

[7] Tillich, M., Lehwark, P., Pellizzer, T., Ulbricht-Jones, E. S., Fischer, A., Bock, R., Greiner, S. (2017). GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic acids research*, 45(W1), W6-W11.

[8] Lohse, M., Drechsel, O., Bock, R. (2007). OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. *Current genetics*, 52(5-6), 267-274.

[9] Beier, S., Thiel, T., Münch, T., Scholz, U., Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics*, 33(16), 2583-2585.

[10] Kurtz, S., Choudhuri, J. V., Ohlebusch, E., Schleiermacher, C., Stoye, J., Giegerich, R. (2001). REPuter: the manifold applications of repeat analysis on a genomic scale. *Nucleic acids research*, 29(22), 4633-4642.

[11] Tamura, K., Stecher, G., Peterson, D., Filipowski, A., Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular biology and evolution*, 30(12), 2725-2729.

[12] Akpınar, B. A., Avsar, B., Lucas, S. J., Budak, H. (2012). Plant abiotic stress signaling. *Plant signaling & behavior*, 7(11), 1450-1455.

[13] Avsar, B., Esmacili Aliabadi, D. (2015). Putative microRNA analysis of the kiwifruit *Actinidia chinensis* through genomic data. *International Journal of Life Sciences Biotechnology and Pharma Research*, 4(2), 96-99.

[14] Avsar, B., Esmacili Aliabadi, D. (2017). In silico analysis of microRNAs in *Spinacia oleracea* genome and transcriptome. *International Journal of Bioscience, Biochemistry and Bioinformatics*, 7(2), 84-92.

[15] Avsar, B., Esmacili Aliabadi, D. (2017). Identification of microRNA elements from genomic data of European hazelnut (*Corylus avellana* L.) and its close relatives. *Plant Omics*, 10(4), 190-196.