DNA BARCODING OF JOJOBA (*Simmondsia chinensis*) PLANTS CULTIVATED IN EGYPT USING *rbcL* GENE

S. M. KHALIL[1*], S. D. IBRAHIM[2], SHAIMAA M. AHMED[3], AND K. S. ABDALLA[3]

1.Department of Plant Genetic Transformation, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center, Giza, Egypt (corresponding author).*

2.Department of Genome Mapping, Agricultural Genetic Engineering Research Institute(AGERI), Agricultural Research Center, Giza, Egypt.

3.Department of Plant Molecular Biology, Agricultural Genetic Engineering Research Institute (AGERI), Agricultural Research Center, Giza, Egypt.

Jojoba (*Simmondsia chinensis*) is a dioecious shrub, native to the southwestern United States and northern Mexico, it is a nontraditional crop that can grow in arid and semi-arid areas, in addition, it is naturally adapted to saline soils and high-temperature environments (Al-Ani *et al*., 1972). Jojoba cultivated mainly for its seed oil, which contains liquid wax very similar to whale sperm in value. Seed oil is highly valued for their use in a wide range of pharmaceutical industry, also as biodegradable lubricants in the motor industry and as a biofuel product (Aburjai and Natsheh 2003; Baldwin, 1988). The use of jojoba in biofuel production has many advantages because the plant is non-edible, unlike corn or sugarcane, and can be irrigated with low quality water and cultivated in areas with high levels of drought, salinity and heat. The nature of the environment and weather in Egypt is suitable and encouraging for jojoba plant (Borlaug, 1985). Indeed, one of the main corners of the Egyptian economy is Agriculture; therefore, improving agricultural systems and irrigation strategies must be of great concern. Because of the above-mentioned reasons, the increase of jojoba cultivation is recommended. The identification and differentiation of male and female jojoba plants are very crucial during cultivation. Therefore, the use of traditional methods based on morphological criteria as well as the molecular methods is very important approaches. DNA barcoding is a tool for species identification that amplify a specific region of DNA and sequencing to create a global database of living organisms (Hebert *et al*., 2003). CBOL Plant Working Group (Consortium for the Barcode of Life: www.barcoding.si.edu) proposed portions of two coding regions from chloroplast genome (*rbcL* and *matK*) as a "core barcode" for plants in 2009 and to be supplemented with additional regions as required. In recent years DNA barcoding has been developing rapidly for a wide range of applications. Four plant DNA barcode markers, *rbcL* (Ribulose bisphosphate carboxylase), *matK* (MaturaseK),

Egypt. J. Genet. Cytol., 49:245-254, July, 2020
Web Site (*www.esg.net.eg*)

trnH-psbA (intergenic spacer region) and *ITS2* (internal transcribed spacer 2), have been developed and used. Most plant DNA barcodes are located in the chloroplast genome, either within coding sequences (such as *rbcL* and *matK*) or in intergenic regions (such as *trnH-psbA*) (Group, *et al*., 2009). However, some nuclear loci have been used as DNA barcodes, such as the internal transcribed spacer of the ribosomal DNA (*ITS*) (Bolson *et al*., 2015). Chase *et al*., (1993) employed *rbcL* as a marker for several early plant phylogenetic studies and soon this marker became popular in that field. The sequence analysis resulted from the studies of *rbcL* gene marker produced a large sequence database. Therefore, *rbcL* became an attractive candidate for molecular identification of plant species. Our aim here is to identify jojoba plants using *rbcL* DNA barcode.

MATERIALS AND METHODS

Plant materials

The analysis involved a number of one male and fifteen female plants (clones) growing at Go-green for Agricultural Investment and Development Company farm, Abo-Ghaleb, Giza, Egypt (GIADC).

DNA extraction, amplification and sequencing

Total DNA was extracted from fresh leaves using DNeasy Plant Mini Kit (QIAGEN, Germany) according to the manufacturer's instructions. After purification of the DNA, *rbcL* DNA barcode region of chloroplast DNA was amplified *via* polymerase chain reactions (PCR). Primers used in the amplifications were

rbcLa-F:(5'ATGTCACCACAAACAGAGACTAAAGC-3') and rbcL724-R: (5'-TCGCATGTACCTGCAGTAGC-3').

For each PCR reaction, 1 µL (25 ng) of total DNA was included in a 50 µL reaction mixture containing 2 units Taq DNA polymerase (GoTaq, Promega), 1X buffer, 20 µM each primer, 0.2 mM dNTPs and ultra-pure water to a final volume of 50 µL. Amplification of the target region was conducted with a thermal cycler (Applied Biosystems, USA) under the following conditions: initial denaturation at 95℃ for 2 min, 35 cycles of 95ºC for 30 s, 55℃ for 30 s, and 72ºC for 1 min, followed by a final extension at 72ºC for 5 min. PCR products were examined on a 1.5% agarose gel in 1× TBE buffer, then visualized on UV light and photographed using a Gel Documentation System (BIO-RAD 2000, USA). After that the PCR fragments were purified using EZ-10 spin column (Biobasic, USA) for the purification of PCR products. The purified fragments were sent to Macrogen Europe B.V. company (Netherlands) for sequencing.

Data analysis

The obtained sequence data were adjusted manually as needed for each sequence. The sequence data generated in this study were analyzed separately in the

GenBank (www. https://www.ncbi.nlm.nih.gov) for similarity search using Nucleotide BLAST database. The *rbcL* sequences were matched with the query sequences and available *rbcL* sequences of the examined plant species available in the GenBank. After that DNA sequences were aligned using multiple sequence alignment by MAFFT (https://www.genome.jp/tools-bin/mafft). All sequences were adjusted manually and submitted to GenBank, USA. Phylogenetic analyses were conducted with the same database. The phylogenetic trees were inferred with the maximum likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993).

RESULTS AND DISCUSSION

Cultivation of jojoba plants have spread in many areas of the world (Benzioni *et al*., 2005) including the Middle East. Few studies have been made on comparative analysis at the molecular genetic level to detect genetic distances among genotypes existing in different regions. Therefore, these studies for jojoba plants are an important task. This study was performed to investigate the possibility of using rbcL DNA barcode for identifying jojoba plants. The amplification of rbcL gene fragment was conducted using rbcLa-F and rbcL724-R primers to amplify approximately 550 bp fragment. As shown in Fig. (1), the amplification of *rbcL* yielded PCR products for all the tested samples. Amplification and sequencing success rate are the most important criteria to evaluate DNA barcoding for plant identification (CBOL Plant Working Group *et al*., 2009). The success rates of amplification and sequencing of *rbcL* fragment were 100%. Our results showed higher universality and success rates similar to those of Ibrahim *et al*., (2019) who conducted rbcL barcode for the identification of different Quinoa genotypes. While, Pei *et al*., (2015) performed *rbcL* barcode with 90%-100% success rate in the forest plants. While other studies showed a success rates up to 90% (Kang *et al*., 2017; Huang *et al*., 2015) for *rbcL* in the tropical forest plants. This study confirms that *rbcL* universality as DNA barcode could be obtained with *rbcL* primers from a wide range of plant species (e.g., Lahaye *et al*., 2008; Parmentier *et al*., 2013). Different candidate gene regions were used as potential barcodes for plants including *rbcL*, *matK*, *trnH-psbA* and *ITS2* (Kress & Erickson, 2007; Taberlet *et al*., 2007). However, the most used are *rbcL* and *matK* for plant DNA barcode. In our study we have used *rbcL* because many authors indicated that the success rate of identification is more than *matK* (Amandita *et al*., 20019; Kang *et al*., 2017; Khidir and Lawrence 1999). Also the use of *matK* as a barcode has been less interest mainly because universal primers are not available (Bafeel *et al*., 2011; Dong *et al*., 2015). Furthermore, Hollingsworth *et al*., (2011) indicated that *matK* still needs optimization in regard to primer combinations and needs to be adapted to specific taxonomic groups. The obtained sequences of the 16 jojoba samples were matched with the reference sequences in BLASTn (Altschul

et al., 1990). The query sequences were identified up to species level with 98 or 100% in either of the algorithms for all samples. Table (1) summarize the obtained results. Samples 1-12 and 16 have the same identity 99 and 100% with *rbcL* gene from different jojoba plants submitted in GenBank. While samples No. 13 and 14 produced 99 % and 98-99% similarity, respectively. On the other hand, sample No. 15 produced 98-99% similarity. Our results are consistent with those obtained by Ghareb *et al*., (2020) who carried out DNA barcoding using *rbcL* and *matK* genes for *Phlomis aurea* plant. The drawing of a phylogenetic tree represents the best hypothesis about how a set of species or other groups evolved from a common ancestor and how they are related to each other. Based on multiple sequence alignment, a phylogenetic relationship of tested samples using *rbcL* was constructed. As shown in Fig. (2), the pattern of branching in a phylogenetic tree reflects how species or other groups evolved from a series of common ancestors. There are two main branches from the common ancestor for all the tested clones. The produced pattern indicated that samples1-12 and 16 have the same cluster with the same branch and very related to each other. Also sample 13 is related to the same group but with a different branch with the same branch point. The other branch containing samples 14 and 15 which are closely related to each other but less related to the other samples. These results are expected due to the tested samples were jojoba clone plants from the same cultivated area as indicated before. As shown in Table (2), the obtained DNA sequence data for all samples were submitted to GenBank for providing the accession numbers for the nucleotide sequences. Definitely the increase of using plant DNA barcode will enrich the database with new sequences which will establish more the use of this technique in a wide range for the identification of different plants. In conclusion, our results support the claim that DNA barcoding in general can provide fast and reliable species identification, especially for the economically important plants.

SUMMARY

Precise identification of jojoba plant species using DNA barcode is very important because of the economic value of its seed oil. In this study, we selected 16 jojoba trees (clones) including one male and 15 female trees cultivated in the same farm to conduct this experiment. DNA samples were extracted from leaf tissues and subjected to PCR amplification using specific primers for *rbcL* gene fragment. The amplification was successful for all tested samples producing the same fragment with a size about 550 bp. Sequence analysis was performed for the purified products and consequently subjected to GenBank database analysis. The data analysis produced by BLASTn database revealed that the similarity search was from 98-100% with *rbcL* gene region form available DNA sequence in GenBank. Multiple sequence alignment was performed after confirmation of the DNA

sequence. The alignment clearly indicated the high similarity between all DNA sequences for the tested samples. Furthermore, a phylogenetic tree was produced from the alignment process which revealed the close relationship between the tested samples. The tree can be divided into two main branches one containing samples 1-13 and 16 while the other branch contains samples 14 and 15. These results indicate that the *rbcL* region is a good option for molecular identification of jojoba species. Finally, the obtained sequence data were submitted to GenBank to provide the accession numbers for the nucleotide sequences.

ACKNOWLEDGEMENTS

We would like to thank Go-green for Agricultural Investment and Development Company (GIADC) Giza, Egypt, for funding this research and providing the plant materials.

REFERENCES

Aburjai T. and Natsheh F. M., (2003). Plants used in cosmetics. Phytother. Res., 17: 987-1000.

Al-Ani H. A., Strain B. R. and Mooney H. A., (1972). The physiological ecology of diverse populations of the desert shrub *Simmondsia chinensis*. Journal of Ecology, 60: 41-57.

Altschul S. F., Gish W., Miller W., Myers E. W. and Lipman D. J., (1990). Basic local alignment search tool. J. Mol. Biol., 215: 403-410.

Amandita F. Y., Rembold K., Vornam B., Rahayu S., Siregar I. Z., Kreft H. and Finkeldey R., (2019). DNA barcoding of flowering plants in Sumatra, Indonesia. Ecology and evolution, 9: 1858-1868.

Bafeel S. O., Arif I. A., Bakir M. A., Khan H. A., Al Farhan A. H., and Thomas J., (2011). Comparative evaluation of PCR success with universal primers of maturase K (*matK*) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (*rbcL*) for barcoding of some arid plants. Plant Omics, 4: 195-198.

Baldwin A. R., (1988). Proceedings: Seventh International Conference on Jojoba and Its Uses. In 7th International Conference on Jojoba and Its Uses, Phoenix, Ariz. (USA), 1988. American Oil Chemists' Society. p: 318-342.

Benzioni A., Mills D., Van Boven M. and Cokelaere M., (2005). Effect of genotype and environment on the concentration of simmondsin and its derivatives in jojoba seeds and Foliage-Industrial Crops and Products, 21: 241-249.

Bolson M., Smidt E. D. C., Brotto M. L. and Silva-Pereira V., (2015). *ITS* and *trnH-psbA* as efficient DNA barcodes to identify threatened commercial woody angiosperms from southern Brazilian Atlantic

rainforests. PloS one, 10 (12): e0143049.

Borlaug N. E., (1985). Jojoba: New Crop for Arid Lands, New Raw Material for Industry, The National Academies Press Vol. 53.

Chase M. W., Soltis D. E., Olmstead R. G., Morgan D., Les, D. H., Mishler B. D., .and Kron K. A., (1993). Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Annals of the Missouri Botanical Garden, 528-580.

CBOL Plant Working Group (2009). A DNA barcode for land plants. Proc. Natl. Acad. Sci., USA 106:12794-12797.

Dong W., Xu C., Li C., Sun J., Zuo Y., Shi S., and Zhou S., (2015). ycf1, the most promising plastid DNA barcode of land plants. Scientific Reports, 5: 8348.

Ghareb H. E., Ibrahim S. D. and Hegazi G. A., (2020). Micropropagtion and DNA-barcoding of the endangered endemic *Phlomis aurea* plant of Saint Katherine. PO J., 13: 65-77.

Hebert P. D., Cywinska A., Ball S. L. and Dewaard J. R., (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270 (1512): 313-321.

Hollingsworth P. M., Graham S. W. and Little D. P., (2011). Choosing and using a plant DNA barcode. PloS one, 6 (5): e19254.

Huang X., Ci X., John G. and Li J., (2015). Application of DNA Barcodes in Asian Tropical Trees - A Case Study from *Xishuangbanna*. Nature Reserve, Southwest China. PLoS one 10: e0129295.

Ibrahim S. D., Abd El-Hakim A. F., Ali H. E., Abd El-Maksoud R. M., (2019) Genetic differentiation using ISSR, SCoT and DNA Barcoding for Quinoa genotypes. Arab J. Biotech., 22: 103-118.

Kang Y., Deng Z., Zang R. and Long W., (2017). DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. Scientific reports, 7: 1-9.

Khidir W. H. and Lawrence, A. A., (1999). Evolutionary implications of *matK* indels in *Poaceae*. Am. J. Bot., 86: 1735-1741.

Kress W. J. and Erickson D. L., (2007). A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-*

psbA spacer region. PLoS one, 2: e508.

Lahaye R., van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G. and Savolainen, P., (2008). DNA barcoding the floras of biodiversity hotspots. Proceedings of the National Academy of Sciences, 105: 2923-2928.

Parmentier I., Duminil J., Kuzmina M., Philippe M., Thomas D. W., Kenfack D., and Hardy O. J., (2013). How effective are DNA barcodes in the identification of African rainforest trees? PLoS ONE, 8: e54921.

Pei N., Erickson D. L., Chen B., Ge X., Mi X., Swenson N. G., and Hao Z., (2015). Closely-related taxa influence woody species discrimination via DNA barcoding: evidence from global forest dynamics plots. Scientific Reports, 5: 1-9.

Taberlet P., Coissac E., Pompanon F., Gielly L., Miquel C., Valentini A., and Willerslev E., (2007). Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Research, 35: e14.

Tamura K. and Nei M., (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol., 10: 512-526.

Table (1): The homologous sequences best matching the *rbcL* sequences based on the BLASTn analysis.

Description	Accession No.	Idententity %	E value
GIADC_1 -12 and 16			
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	NC_040935.1	100.00%	0
Simmondsia chinensis voucher Yi14353 (KUN) plastid, complete genome	MK397929.1	100.00%	0
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	MK397898.1	100.00%	0
Simmondsia chinensis voucher Hosam00080 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) gene, partial cds; chloroplast	JX571893.1	100.00%	0
Simmondsia chinensis ribulose-1,5-bisphosphate carboxylase oxygenase, large subunit (*rbcL*) gene, partial cds; chloroplast gene for chloroplast product	AF093732.1	99.39%	0
GIADC_13			
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	NC_040935.1	99.79%	0
Simmondsia chinensis voucher Yi14353 (KUN) plastid, complete genome	MK397929.1	99.79%	0
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	MK397898.1	99.79%	0
Simmondsia chinensis voucher Hosam00080 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) gene, partial cds; chloroplast	JX571893.1	99.79%	0
Simmondsia chinensis ribulose-1,5-bisphosphate carboxylase oxygenase, large subunit (*rbcL*) gene, partial cds; chloroplast gene for chloroplast product	AF093732.1	99.58%	0
GIADC_14			
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	NC_040935.1	99.60%	0
Simmondsia chinensis voucher Yi14353 (KUN) plastid, complete genome	MK397929.1	99.60%	0
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	MK397898.1	99.60%	0
Simmondsia chinensis voucher Hosam00080 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) gene, partial cds; chloroplast	JX571893.1	99.60%	0
Simmondsia chinensis ribulose-1,5-bisphosphate carboxylase oxygenase, large subunit (*rbcL*) gene, partial cds; chloroplast gene for chloroplast product	AF093732.1	98.99%	0

Table (1): Cont''

GIADC_15			
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	NC_040935.1	99.80%	0
Simmondsia chinensis voucher Yi14353 (KUN) plastid, complete genome	MK397929.1	99.80%	0
Simmondsia chinensis voucher 14CS9341 (KUN) plastid, complete genome	MK397898.1	99.80%	0
Simmondsia chinensis voucher Hosam00080 ribulose-1,5-bisphosphate carboxylase/oxygenase large subunit (*rbcL*) gene, partial cds; chloroplast	JX571893.1	99.80%	0
Simmondsia chinensis ribulose-1,5-bisphosphate carboxylase oxygenase, large subunit (*rbcL*) gene, partial cds; chloroplast gene for chloroplast product	AF093732.1	99.19%	0

Table (2): Sample I.D. and Accession numbers provided by GenBank after submission of the DNA sequences.

Sample I.D.	Accession No.
GIADC_1	MT895755
GIADC_2	MT895758
GIADC_3	MT895757
GIADC_4	MT895756
GIADC_5	MT895746
GIADC_6	MT895747
GIADC_7	MT895749
GIADC_8	MT895745
GIADC_9	MT895754
GIADC_10	MT895751
GIADC_11	MT895753
GIADC_12	MT895752
GIADC_13	MT895759
GIADC_14	MT895744
GIADC_15	MT895748
GIADC_16	MT895750

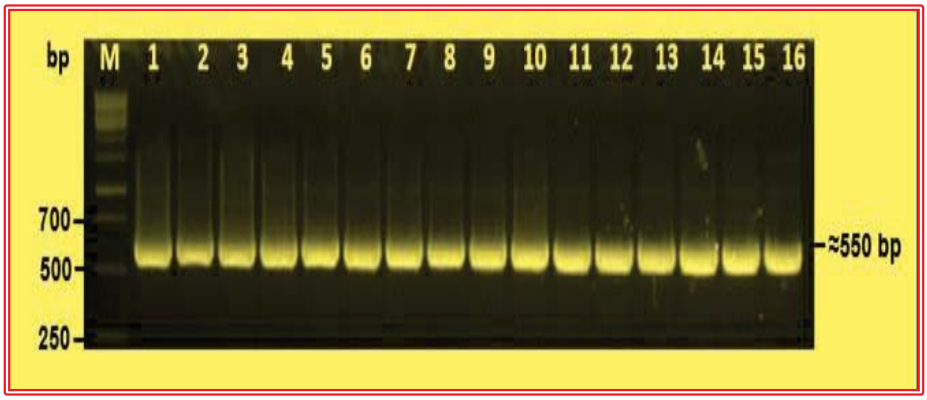

Fig.(1): PCR amplification of *rbcL* gene fragment for different jojoba plants. M: 1kb DNA ladder Marker. Lanes 1-16: Amplified fragments of *rbcL* gene for the tested samples.

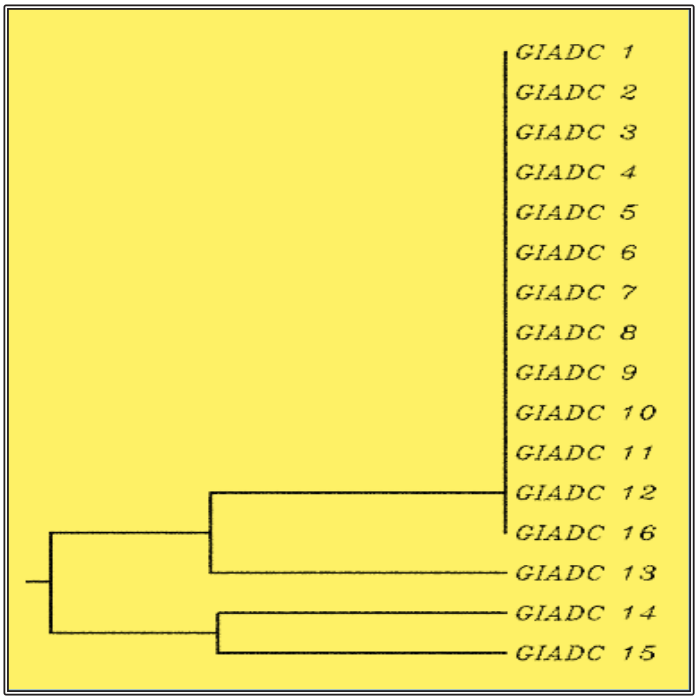

Fig. (2): Phylogenetic tree constructed by Neighbor joining method of the samples representing the selected jojoba plants based on *rbcL* gene sequences.