MORPHOLOGICAL AND MOLECULAR CHARACTERIZATION OF *Citrus* GERMPLASM IN SRI LANKA

W. A. N. R. Wickramaarachchi[1*], W. W. P. Rodrigo[2] and S. R. Weerakoon[1]
[1]Department of Botany, [2]Department of Zoology, The Open University of Sri Lanka, Sri Lanka

Sri Lanka has a rich variety of *Citrus* species, ranging from wild-growing limes in village backyards to cultivated hybrid mandarins and oranges in commercial plantations. Despite their economic and cultural importance, the true diversity of *Citrus* in Sri Lanka has not been fully explored or documented. The main objective of this project is to fill that gap by identifying and characterizing *Citrus* species across the island using both traditional morphological methods and modern molecular tools. Morphological characterization involved the documentation of traits such as fruit, leaf, tree and floral morphology. Quantitative data were scandalized using z- score normalization, while qualitative data were encoded numerically. Species distances were calculated using Gower distance in the R programming environment. A dendrogram was constructed using the word's method of hierarchical clustering in IBM SPSS. Based on the morphological characters, five main clusters were identified. Cluster I - Yak dehi, Key lime, Nas naran, and Kaffir lime. Cluster II - Lemon- round, Lemon – long, Cluster III - Sour orange, Sweet orange Cluster IV - Sidaran, Cluster V – Pumelo.
Molecular characterization was carried out through DNA barcoding of chloroplast genes. Genomic DNA was extracted using a modified CTAB method, and PCR amplification was performed using *rbc*L and *mat*K gene markers. PCR products were sequenced at Macrogen, South Korea. The sequences were analyzed using BioEdit software and deposited in GenBank. Phylogenetic trees were constructed using MEGA software to infer genetic relationships among species. The phylogenetic tree based on *mat*K sequences revealed three major clusters: Cluster I - Sour orange, sweet orange, Nas naran, Yak dehi, and Pumelo. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime. Cluster III - Sidaran. The phylogenetic tree based on *rbc*L sequences formed two primary clusters. Cluster I - Sour orange, Sweet orange, Nas naran, Yak dehi, Pumelo, and Sidaran. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime. Phylogenetic analysis using both *rbc*L and *mat*K markers revealed consistent clustering patterns, with the only notable exception being the placement of Sidaran. These findings provide a foundational understanding of the morphological and genetic diversity of *Citrus* species in Sri Lanka and contribute to future conservation and breeding programs.

Keywords: Citrus diversity, Sri Lanka, molecular characterization, morphological characterization

[*]*Corresponding Author: wanr9556443@gmail.com*

MORPHOLOGICAL AND MOLECULAR CHARACTERIZATION OF *Citrus* GERMPLASM IN SRI LANKA

W. A. N. R. Wickramaarachchi[1*], W. W. P. Rodrigo[2] and S. R. Weerakoon[1]
[1]Department of Botany, [2]Department of Zoology, The Open University of Sri Lanka, Sri Lanka

INTRODUCTION

Citrus is one of the world's most economically and medicinally important plant genera which belongs to the subfamily Aurantioideae (Citroideae) of the family Rutacea. It is believed that *Citrus* was originated in Southeast Asia (Calabrese, 1992) and was cultivated in China in 2500BC (Nicolosi et al., 2000).

Citrus possess many essential phytochemical molecules like ferulic acid, hydrocinnamic acid, cyanidin glycoside, hesperidin, naringin, polymethoxalated flavone, vitamin C, and carotenoids (Ma et al., 2020), which have antioxidant, anti-inflammatory, anti-cancer, anti-tumor, anti-fungal, anti-bacterial, and blood clot inhibition activities (Abesinghe et al., 2007).

Sri Lanka is very famous for superior quality *Citrus* species from ancient times and this fruit is being used as fresh fruits, fresh juices, and in *Citrus* based industries as well as in Ayurveda and traditional system of medicines for treatment of an array of aliments. Recent studies on Sri Lankan *Citrus* species revealed that they have extraordinary contents of phytochemicals and vitamins C as well as they have industrially desirable morphological and biochemical traits (Herath et al., 2015). Therefore, there is a high potential of commercialization of these species as industrial crops as well as pharmaceutical crops in Sri Lanka. However, despite the country's rich *Citrus* diversity, there exists a notable gap in scientific understanding and documentation of *Citrus* species found in Sri Lanka. The absence of comprehensive morphological and molecular characterization limits the potential for informed decision making in *Citrus* cultivation, breeding, and conservation efforts.

Morphological traits such as fruit size, shape, color, leaf characteristics, and tree morphology provide valuable insight into the diversity and taxonomy of *Citrus* species. By systematic characterization and documentation of these traits, can establish a foundation for accurate species identification, classification, germplasm conservation, and incorporation of breeding programs aimed at developing improved *Citrus* varieties.

Advances in molecular biology offer powerful tools like DNA barcoding for evaluating genetic diversity, phylogenetic relationships, and evolutionary history of *Citrus* species. By analyzing specific DNA regions unique to *Citrus* species, DNA barcodes can be established that can serve as molecular fingerprints, enabling rapid

and accurate species identification, assessment of genetic diversity and phylogenetic relationships among *Citrus* species in Sri Lanka. Molecular characterization of *Citrus* species, using chloroplast DNA markers *mat*K and *rbc*L will enhance species identification, improve phylogenetic resolution, and aid in the prevention of bio piracy.

This research contributes to species identification and germplasm conservation of *Citrus* species available in Sri Lanka.

METHODOLOGY

Samples were collected from Agriculture Research Station, Maduruketiya, Monaragala and home gardens of Kurunegala, Gampaha and Colombo Districts in Sri Lanka.

Morphological Characterization

Seventy-four (74) morphological characters of *Citrus* species were examined and documented using ten replicates per each species, including twelve (12) tree morphological traits, fifteen (15) flower morphological traits, sixteen (15) leaf morphological traits, twenty-four (24) fruit morphological traits, and eight (8) seed morphological traits according to the International Plant Genetic Resource Institute descriptors for *Citrus* (IPGRI,1999).

The data were submitted to analysis of variance (ANOVA) followed by post-hoc separation of means which was done using Fischer's Protected Least Significant Difference (LSD) test at 5% probability (Masuka, Goss & Mazarura, 2012), using SPSS statistical software (SPSS Inc., Chicago). Quantitative data were analyzed using z- score normalization and qualitative data were encoded with numerical value. Distances between species were calculated using Gower distance in R program for mixed data (Ricaut et al., 2010). Dendrogram was constructed using ward's method of hierarchical clustering in IBM- SPSS analytical software (Sosa & Kranzfelder, 2017).

Molecular characterization

DNA Extraction and Visualizing

DNA was isolated from young leaves of *Citrus* species according to the Healey et al. (2014) specification with some modifications. DNA extraction was done by grinding young *Citrus* leaves with pre-heated CTAB extraction buffer [2% CTAB, 100 mM Tris-HCl, 2.5 mM EDTA, 1.5 mM NaCl, 1% PVP], followed by 1-hour incubation of samples at 65 ºC. Samples which were centrifuged at 13,000 g for 10 minutes. Supernatant was transferred into a new tube, an equal volume of chloroform:iso-amylalcohol (24:1) was added and centrifuged at 13,000 g for 10

minutes. This step was repeated three times. The upper aqueous layer was transferred to a new tube and equal volume of cold (-20 ºC) isopropanol and 4 M NaCl was added. Samples were incubated at 1 hour at -20 ºC and centrifuged at 16,000 rpm for 15 minutes at 4 ºC. The DNA pellets were washed with 70% ethanol followed by 100% ethanol. Air dried DNA pellets were re-suspended in Nuclease free water and stored at -20 ºC for future usage.

The volume of 3 µL genomic DNA from each *Citrus* species was mixed with the 2 µl of loading dye and the mixture was loaded into 0.8% agarose gel. The electrophoresis was carried out for 45 minutes at 80 V and it was visualized using a gel visualizer system (Khan et al., 2009). The quality and quantity of the undiluted and ×100 diluted genomic DNA samples were measured by Nano-drop method (Nakayama et al., 2016; Simbolo et al., 2013).

Amplification of *rbcL* Gene by Polymerase Chain Reaction (PCR)

PCR were carried out according to the protocol of Gori et al. (2019). PCR were performed using ~50 ng of genomic DNA in a 15 µl of reaction. PCR cocktail containing 7.5 µl of 2X Promega Gotaq master mix (2X Gotaq reaction buffer (pH 8.5), 400 µM dATP, 400 µM dGTP, 400 µM dCTP, 400 µM dTTP and 3 mM $MgCl_2$), 0.5 µl of each 5 µM primers, and 5.5 µl of deionized water. The primer pairs of *rbc*L_F1 (5‟- ATG TCA CCA CAA ACA GAG ACT AAA GC- 3‟) and rbcL_R (5‟- GAA ACG GTC TCT CCA ACG CAT -3‟) were used as forward and reverse primers respectively. The reaction mixture was mixed, spinning at 6,500 rpm for 1 minute. The amplification was carried out in a thermal cycler (primus 96). PCR condition was as follows: denaturation at 94 ºC for 5 minutes, then 35 cycles at 94 ºC 30 seconds followed by 30 seconds at a temperature 62 ºC and extension cycle at 72 ºC for 1 minute, a final extension was performed at 72 ºC for 8 minutes followed by stabling temperature at 4 ºC for infinity.

Amplification of *mat*K Gene by Polymerase Chain Reaction (PCR)

PCR were carried out according to the protocol of Gori et al. (2019). PCR were performed using ~50 ng of genomic DNA in a 15 µl of reaction PCR cocktail containing 7.5 µl of 2X Promega Gotaq master mix (2X Gotaq reaction buffer (pH 8.5), 400 µM dATP, 400 µM dGTP, 400 µM dCTP, 400 µM dTTP and 3 mM $MgCl_2$), 0.5 µl of each 5 µM primers, and 5.5 µl of deionized water. The primer pairs of KIM_3F (5‟- CGT ACA GTA CTT TTG TGT TTA CGA G- 3‟) and KIM_1R (5‟- ACC CAG TCC ATC TGG AAA TCT TGG TTC -3‟) were used as forward and reverse primers respectively. The reaction mixture was mixed, spinning at 6,500 rpm for 1 minute. The amplification was carried out in a thermal cycler (primus 96). PCR conditions were as follows: denaturation at 94 ºC for 5 minutes, then 30 cycles at 94 ºC for 30 seconds followed by 30 seconds at a temperature of 52 ºC and extension cycle at 72 ºC for 1 minute, a final extension will be performed at 72 ºC for 8 minutes followed by stabling temperature at 4 ºC for infinity.

Visualization of Amplified PCR products and Sequencing the Amplified PCR Products

A volume of 7.0 µL of amplified products were mixed with 2 µl of loading dye and loaded into 1.5% agarose gel. A volume of 4.5 µl of 100 base pair (bp) ladder was used as a marker. The electrophoresis was carried out for 1 hour 15 minutes at 80 V and visualized using a gel documentation system. The PCR products of *Citrus* were sent to Macrogen, South Korea to obtain the sequences. The sequences were analyzed using BIOEDIT (Sevġndġk & Yalçin, 2018) program. Sequences were deposited in GenBank. The phylogenic tree was constructed using MEGA software (Amar, 2012) to identify genetic relationships between the species. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model with thousand bootstrap replicates. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. This analysis involved 10 nucleotide sequences. Codon positions included were 1st+2nd+3rd+Noncoding.

RESULTS AND DISCUSSION

Dendrogram constructed using morphological data has five main clusters. Cluster I - Yak dehi, Key lime, Nas naran, Kaffir lime. Cluster II - Lemon- round, Lemon – long, Cluster III - Sour orange, Sweet orange, Cluster IV - Sidaran, Cluster V – Pumelo.

However, the phylogenetic tree based on *mat*K DNA sequences revealed three main clusters: Cluster I - Sour orange, Sweet orange, Nas naran, Yak dehi, and Pumelo. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime. Cluster III – Sidaran.

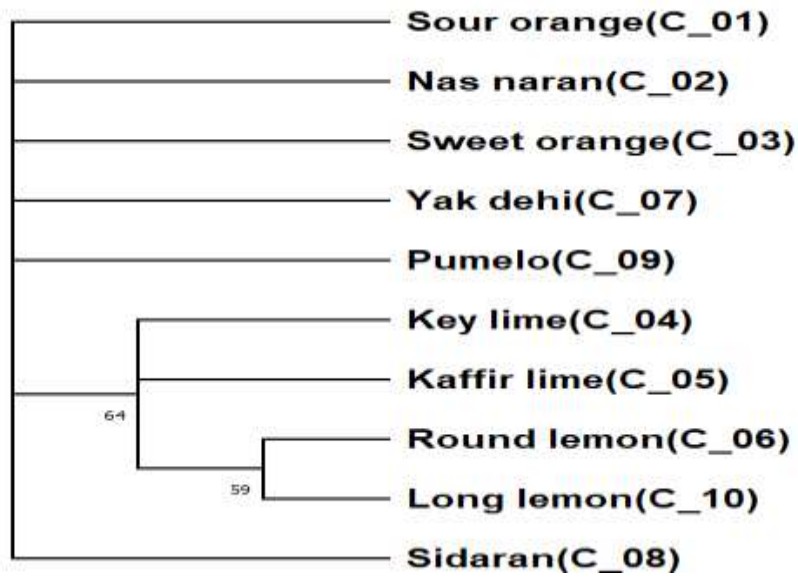

Figure 01. Phylogenetic tree based on *mat*K DNA sequences

In contrast, the phylogenetic tree based on *rbc*L DNA sequences formed two main clusters. Cluster I - Sour orange, Sweet orange, Nas naran, Yak dehi, Pumelo, and Sidaran. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime.

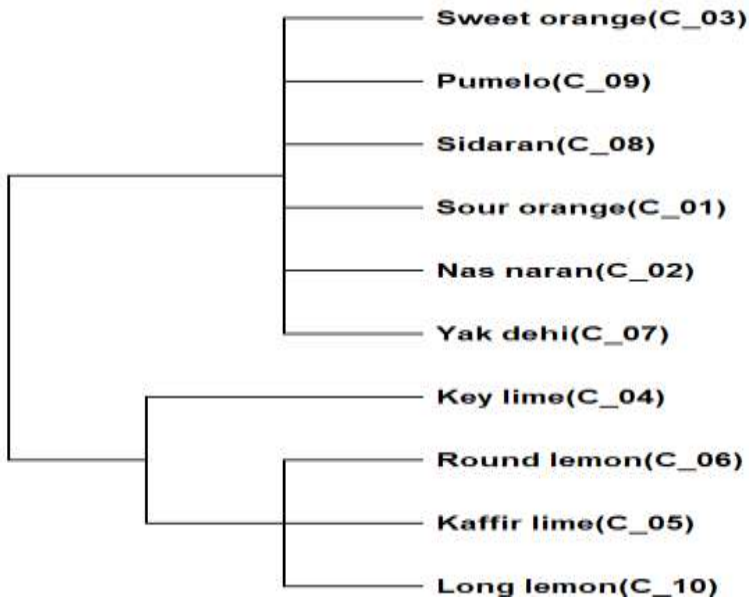

Figure 02. Phylogenetic tree constructed based on *rbc*L gene sequences

The phylogenetic analysis using both *rbc*L and *mat*K markers revealed a consistent clustering pattern, with Cluster I and Cluster II showing similar grouping across both markers. The only major exception was Sidaran, which formed a separate cluster (Cluster III) in the *mat*K-based tree but was grouped under Cluster I in the *rbc*L-based tree. This overall consistency in clustering highlights the reliability of these chloroplast DNA markers in the classification of *Citrus* species.

CONCLUSIONS/RECOMMENDATIONS

The study on Morphological and Molecular Characterization of *Citrus* germplasm in Sri Lanka revealed the following findings.

Through morphological analysis, identified and tested Sri Lankan *Citrus* germplasm can be clustered in to five main clusters. Cluster I - Yak dehi, Key lime, Nas naran, Kaffir lime. Cluster II - Lemon- round, Lemon – long, Cluster III - Sour orange, Sweet orange, Cluster IV - Sidaran, Cluster V – Pumelo.

The phylogenetic tree based on *mat*K DNA sequences revealed three main clusters: Cluster I - Sour orange, Sweet orange, Nas naran, Yak dehi, and Pumelo. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime. Cluster III - Sidaran.

The phylogenetic tree based on *rbc*L DNA sequences formed two main clusters. Cluster I - Sour orange, Sweet orange, Nas naran, Yak dehi, Pumelo, and Sidaran. Cluster II - Long lemon, Round lemon, Key lime, and Kaffir lime. The phylogenetic analysis using both *rbc*L and *mat*K markers revealed a consistent clustering pattern, with Cluster I and Cluster II showing similar grouping across both markers.

It is recommended that molecular characterization needs to secure a barcode of a further two to three marker gene regions to obtain clear ideal clusters.

REFERENCES

Abeysinghe, D. C., Li, X., Sun, C., Zhang, W., Zhou, C., & Chen, K. (2007). Bioactive compounds and antioxidant capacities in different edible tissues of citrus fruit of four species. Food chemistry, 104(4), 1338-1344.

Calabrese F (1992). The history of citrus in the Mediterranean countries and Europe. Proc Int Soc Citricult 1:35–38.

Federici, C.T., Fang, D.Q., & Scora, R.W. (1998). Phylogenetic relationship within the genus Citrus (Rutaceae) and related genera as revealed by RFLP and RAPD analysis. Theor. Appl. Genet. 96, 812–822.

Gori, M., Pecchioli, S., Giordani, E., Saeedi, M. A., Wafa, F. H., & Biricolti, S. (2019). Barcoding assessment of the Citrus species cultivated in eastern Afghanistan. Advances in Horticultural Science, 33(3), 403-408.

Herath, H. D., Chamikara, M. D., Dissanayake, D. R. R., Dissanayake, M. D. M. I., & Ishan, M. (2016). A comparative assessment of the antibacterial activity in fruit juice of Sri Lankan sweet orange cultivars vis a vis sour orange.

IPGRI- International Plant Genetic Resources Institute. Descriptors for Citrus. Rome: IBPGR (1999).

Ma, G., Zhang, L., Sugiura, M., & Kato, M. (2020). Citrus and health. In *The genus citrus* (pp. 495-511). Woodhead Publishing

Nicolosi, E., Deng, Z. N., Gentile, A., La Malfa, S., Continella, G., & Tribulato, E. (2000). Citrus phylogeny and genetic origin of important species as investigated by molecular markers. *Theoretical and Applied Genetics*, *100*, 1155-1166.

Zhang, Y., Muyrers, J. P., Testa, G., & Stewart, A. F. (2000). DNA cloning by homologous recombination in Escherichia coli. *Nature biotechnology*, *18*(12), 1314-1317.