

The establishment of the species-delimits and varietal-identities of the cultivated germplasm of *Cucumis melo* and *Cucumis sativus* in Sri Lanka using morphometric, organoleptic and phylogenetic approaches

P. R.S. Pathirana

pathiranaprs@gmail.com

Rajarata University of Sri Lanka

E. Y. Fernando

Rajarata University of Sri Lanka

A. M.K.R. Bandara

Rajarata University of Sri Lanka

D. M.D. Dissanayake

Rajarata University of Sri Lanka

A. M.J.B. Adikari

Rajarata University of Sri Lanka

Research Article

Keywords: *Cucumis melo*, Kekiri, underutilized vegetable, germplasm, DNA barcoding, phylogenetics, phenomics

Posted Date: October 23rd, 2025

DOI: <https://doi.org/10.21203/rs.3.rs-7757768/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

The establishment of the species-delimits and varietal-identities of the cultivated germplasm of *Cucumis melo* and *Cucumis sativus* in Sri Lanka using morphometric, organoleptic and phylogenetic approaches.

P.R.S. Pathirana^{1*}, E.Y. Fernando¹, A.M.K.R. Bandara², D.M.D. Dissanayake², & A.M.J.B. Adikari²

¹Faculty of Applied Sciences, Rajarata University of Sri Lanka, Mihintale, Anuradhapura, Sri Lanka

²Faculty of Agriculture, Rajarata University of Sri Lanka, Puliyankulama, Anuradhapura, Sri Lanka

Corresponding author: P.R.S. Pathirana

*Correspondence E-mail: rasikasp2009@gmail.com, pathiranaprs@gmail.com

ORCID: <https://orcid.org/0009-0002-8989-4639>

Abstract

Cucumis melo L. (vern. Kekiri) is an important but underutilized vegetable in Sri Lanka, traditionally used for both culinary and medicinal purposes. Despite its cultural significance, the genetic and phenotypic diversity of Sri Lankan *C. melo* germplasm has remained largely unexplored. In contrast, *C. melo* and *C. sativus* cultivars in developed countries have been extensively characterized, with rich genomic resources available to support crop improvement. This study aimed to establish species delimitation and varietal identities of cultivated *C. melo* and *C. sativus* germplasm in Sri Lanka using morphometric, organoleptic, and molecular approaches. An island-wide germplasm collection was undertaken, including 11 *C. melo* and 9 *C. sativus* genotypes, which were assessed using IPGRI descriptors, field morphological trials, and consumer preference analyses. Organoleptic evaluation revealed complex sensory profiles across cooking, salad, and fruit varieties, with sweetness emerging as a key determinant of preference in fruit types, while bitterness negatively influenced salad varieties. Molecular barcoding using *rbcL* and *trnH-psbA* markers enabled phylogenetic analysis, species delimitation, and divergence dating. Results indicated clear genetic differentiation between *C. melo* and *C. sativus* germplasm, with divergence estimated at ~23 million years, older than previous reports. Concatenated marker analysis improved resolution of varietal relationships, supporting conservation and breeding applications. This integrative study provides the first comprehensive characterization of Sri Lankan *C. melo* germplasm, establishing a foundation for its promotion as a vegetable, medicinal resource, and valuable genetic pool for future crop improvement.

Keywords: *Cucumis melo*, Kekiri, underutilized vegetable, germplasm, DNA barcoding, phylogenetics, phenomics

Introduction

The genus *Cucumis* (family Cucurbitaceae) includes both cultivated and wild species, widely distributed across tropical and subtropical regions (Liu et al., 2004). *Cucumis* species are valued for their fruits, vegetables, and medicinal properties (Lester, 1996). In Sri Lanka, two species of *Cucumis* are reported: *C. sativus* L. (commonly known as cucumber in English and *Pipinha* in Sinhala) and *C. melo* L. (known as *Kekiri* in Sinhala). These species are distinguished by the shape and texture of the fruit, the ovary shape, and the surface pubescence (Dassanayake and Fosberg, 1988; Rajapaksha, 1998). While *C. sativus* has been extensively studied in international research and breeding programs, the taxonomic status of *C. melo* remains complex. According to Liu et al. (2004), *C. melo* is the most genetically diverse species within the Cucurbitaceae family. This species includes a variety of melon types, such as netted musk melon, cantaloupe, honeydew melon, casaba melon, and several oriental pickling and cooking melons (Kirkbride, 1993). Maggs et al. (1999) identified *Kekiri* as a cultivar in their studies. Japanese researchers have collected *Kekiri* seeds from open markets in Kandy, Sri Lanka, and have morpho-genetically characterized them using isozyme markers, as well as created segregating progenies by crossing with other *Cucumis* species (Okubo et al., 1990). Despite this, *C. melo* in Sri Lanka remains understudied and underutilized as a vegetable compared to *C. sativus*.

The genetic diversity and phylogenetic relationships within the genus *Cucumis* have been studied worldwide. For instance, Kocyan et al. (2007) explored the molecular phylogenetics of Cucurbitaceae species, though *C. melo* was excluded from their sampling. Molecular analysis has suggested that the center of origin for *C. melo* lies mainly in Africa, the Middle East, and Southeast Asia (Mliki et al., 2001). However, wild relatives of *C. sativus* and *C. melo* are primarily found in India, other parts of Asia, and Australia, which challenges the African hypothesis for the species' origin. Researchers have suggested that South Asia and Australia should be considered key regions for further investigation into the genetic diversity of *C. melo* and *C. sativus* (Sebastian et al., 2010). Numerous genetic diversity studies on *C. sativus* and *C. melo* have been conducted using isozyme markers (Knerr et al., 1998), microsatellites (Lv et al., 2012), and randomly amplified polymorphic DNA (RAPD) markers (Nakata et al., 2005), yet none have focused on the landraces and farmer varieties of *C. melo* in Sri Lanka.

Phenomics, which involves large-scale phenotyping of plant populations, is an emerging approach for effectively characterizing plant germplasm. Photosynthetic phenomics, which uses chlorophyll fluorescence measurements, is particularly useful for assessing plant photosynthesis. Key parameters such as ΦII (light utilized for photosynthesis), ΦNPQ (light dissipated as heat to prevent damage), and ΦNO (light potentially causing damage) provide a comprehensive picture of photosynthetic efficiency. Additionally, linear electron flow (LEF) reflects energy movement through the chloroplast, while SPAD measures chlorophyll content (Kramer et al., 2004). Using the hand-held MultispeQ device (Kuhlgert et al., 2016), photosynthetic phenomics has been applied to assess *Cucumis melo* L. germplasm for the first time, marking a novel application of these techniques to this species.

Materials and Methods

Sampling Sites

An island-wide germplasm collection was conducted to capture all farmer varieties, cultivars, ecotypes, and wild varieties of *Cucumis melo* L. and *Cucumis sativus* L. in Sri Lanka. The locations of the sample collection sites, including GPS coordinates, are provided in Appendix 1 (*GPS Locations of Island-wide Germplasm Collection*).

Sample Collection

Landraces, farmer varieties, and cultivars of *Cucumis melo* L. were collected from various growing regions across Sri Lanka. The germplasm collection survey was carried out with input from farmers, agricultural instructors, and other relevant stakeholders. All agro-ecological zones in Sri Lanka that offer favorable conditions for the growth and development of *Cucumis* species were visited. For each collection site, GPS coordinates were recorded, and photographic documentation was taken. Additional data collected included ethnobotanical, ethnocultural, and ethno-medicinal information, as well as the agricultural and culinary potential of each genotype.

The morphological and reproductive characteristics of the plants were documented using a structured questionnaire based on the “Descriptors for Melon (*Cucumis melo* L.)” from the International Plant Genetic Resources Institute (IPGRI) and the European Cooperative Program for Plant Genetic Resources (ECPGR) (Diez et al., 2005; Lue et al., 2004). The collected data were recorded in Google Sheets for analysis (*Appendix 1: GPS Locations of Island-wide Germplasm Collection*).

Field Planting

Approximately 20 genotypes were planted in field trials, including 11 *C. melo* genotypes and 9 *C. sativus* genotypes (for comparison) (Table 2). The planting sites included Homagama (WL3, [AEZ] 6.846277, 80.002361), Cheddikulam (DL3, [AEZ] 8.673346, 80.29473), and Tissamaharama (DL5, [AEZ] 6.282261, 81.283851). The experiments were conducted during the growing seasons of 2020 and 2021.

The experimental layout followed a Randomized Complete Block Design (RCBD) with two replicates (beds) for each genotype (Figure 1). The plants were managed according to the cultivation guidelines provided by the Department of Agriculture (DOA) for *C. sativus*.

Table 1: List of collected genotypes (11 *C. melo* genotypes and 9 *C. sativus* genotypes) and their acronyms

Field Code	Species	Common name	Abbreviation
V1	<i>Cucumis stivus</i> L.	Sassy	Sassy
V2	<i>Cucumis stivus</i> L.	Wealthy	Wealthy
V3	<i>Cucumis stivus</i> L.	Fiesty	Fiesty
V4	<i>Cucumis stivus</i> L.	Ajex	Ajex
V5	<i>Cucumis stivus</i> L.	Chandani	Chandani
V6	<i>Cucumis stivus</i> L.	Treasure	Treasure
V7	<i>Cucumis stivus</i> L.	Kalpiti-White	KW
V8	<i>Cucumis stivus</i> L.	LY58	LY58
V9	<i>Cucumis melo</i> L.	Green-Seen-Giant	GSG
V10	<i>Cucumis melo</i> L.	Green-Seen-Small	GSS
V11	<i>Cucumis melo</i> L.	Yellow-Seen-Small	YSS
V12	<i>Cucumis melo</i> L.	White-Seen-Medium	WSM
V13	<i>Cucumis melo</i> L.	Mal-Kekiri	MK
V14	<i>Cucumis melo</i> L.	Deshiya-Kekiri	DK
V15	<i>Cucumis melo</i> L.	Sigari-Kekri	SK
V16	<i>Cucumis stivus</i> L.	Deshiya-Pipinna	DP
V17	<i>Cucumis melo</i> L.	Long Yellow Vegetable	LYV
V18	<i>Cucumis stivus</i> L.	Thiyambara	TH
V19	<i>Cucumis melo</i> L.	Wild spp-1 Gon Kekiri	GK
V20	<i>Cucumis stivus</i> L.	Thiyambara White	THW

V2P1	V2P2	V2P3	V2P4	V5P1	V5P2	V5P3	V5P4
V9P1	V9P2	V9P3	V9P4	V4P1	V4P2	V4P3	V4P4
V18P1	V18P2	V18P3	V18P4	V19P1	V19P2	V19P3	V19P4
V15P1	V15P2	V15P3	V15P4	V10P1	V10P2	V10P3	V10P4
V6P1	V6P2	V6P3	V6P4	V3P1	V3P2	V3P3	V3P4
V1P1	V1P2	V1P3	V1P4	V15P1	V15P2	V15P3	V15P4
V3P1	V3P2	V3P3	V3P4	V7P1	V7P2	V7P3	V7P4
V14P1	V14P2	V14P3	V14P4	V9P1	V9P2	V9P3	V9P4
V17P1	V17P2	V17P3	V17P4	V18P1	V18P2	V18P3	V18P4
V11P1	V11P2	V11P3	V11P4	V12P1	V12P2	V12P3	V12P4
V5P1	V5P2	V5P3	V5P4	V13P1	V13P2	V13P3	V13P4
V20P1	V20P2	V20P3	V20P4	V17P1	V17P2	V17P3	V17P4
V13P1	V13P2	V13P3	V13P4	V8P1	V8P2	V8P3	V8P4
V12P1	V12P2	V12P3	V12P4	V11P1	V11P2	V11P3	V11P4
V16P1	V16P2	V16P3	V16P4	V14P1	V14P2	V14P3	V14P4
V8P1	V8P2	V8P3	V8P4	V6P1	V6P2	V6P3	V6P4
V7P1	V7P2	V7P3	V7P4	V16P1	V16P2	V16P3	V16P4
V19P1	V19P2	V19P3	V19P4	V20P1	V20P2	V20P3	V20P4
V4P1	V4P2	V4P3	V4P4	V1P1	V1P2	V1P3	V1P4
V10P1	V10P2	V10P3	V10P4	V2P1	V2P2	V2P3	V2P4
B1 (R1)				B2 (R2)			

Figure 1: Field planting layout. 19 Genotypes (G), two replicates (B or R), and four plants (P) per replicate are shown here. All (11 *C. melo* and eight *C. sativus*) genotypes were assessed together.

Morphological Data Collection and Statistical Analysis

The morphological and reproductive characteristics of the plants were documented by growing them in controlled experiments, following the “Descriptors for Melon (*Cucumis melo* L.)” from the International Plant Genetic Resources Institute (IPGRI) and the European Cooperative Program for Plant Genetic Resources (ECPGR) (Diez et al., 2005; Lue et al., 2004) (*Appendix 2: Morphological Descriptors of Melon from IPGRI and ECPGR*).

Both qualitative and quantitative data were collected and subjected to appropriate statistical analyses. The quantitative data were analyzed using multiple regression and cluster analysis techniques, along with dimension reduction through Principal Component Analysis (PCA). Non-parametric data, resulting from the organoleptic assessment, were analyzed using association analysis and PCA to visualize the graphical diversity structure that depicts consumer preferences (Kumari et al., 2019).

Assessment of Organoleptic Properties

The evaluation of consumer preferences for *Cucumis sativus* L. and *Cucumis melo* L. was conducted using three different recipes, selected based on the most popular consumer patterns in Sri Lanka. For the evaluation of *Cucumis* vegetable varieties, preferences for mature fruits and their culinary preparations were assessed according to the procedures outlined in Kumari et al. (2019).

Kekiri drupes from each genotype were harvested at the appropriate maturity stage. The dishes were prepared for 50 panelists according to the recipe provided in Table 1. To prepare the drupes, a small slice from the top of each drupe was peeled off, and the cutting edge was rubbed for 1 minute to remove bitterness. The drupes were then peeled, cut into 2 cm pieces, and the seeds were removed. The dishes were prepared based on the ingredients and quantities listed in Table 3.

A modified sensory evaluation procedure, as described in Kumari et al. (2019), was used for this study. Fifty panelists were invited to rank the desired levels of color, aroma, texture, and overall taste using a three-tier ranking system (1: least preferred, 2: moderately preferred, 3: most preferred). The relative bitterness level was ranked separately using a three-tier system (1: least bitter, 2: moderately bitter, 3: most bitter). Panelists were provided with drinking water between tasting different dishes to avoid any carry-over effects on their taste buds.

Table 2: Ingredients of prepared *Cucumis melo* L. and *Cucumis sativus* L. dishes

Ingredients for 800g of cucumis spp.	Quantity of ingredients
Fenugreek seeds	¼ tsp
Black pepper	½ tsp
Curry powder	1tsp
Turmeric powder	½ tsp
Mustard seeds powder	¼ tsp
Chili powder	1 tsp
Coconut milk	200 ml
Curry leaves	
Green chilies	4
Tamarind seeds	5
Cinnamon	1 inch piece
Garlic cloves	5
Red onion	1
Salt	1 tsp
Vinegar	3ml

Half a teaspoon of mustard seed powder was mixed with 3 ml of vinegar in a separate cup to reduce the bitterness of the mustard. Then, ½ teaspoon of salt was added to the mustard-vinegar mixture, and it was set aside for 30 minutes. Following the traditional *Cucumis melo* cooking method, a small slice was removed from the top of the fruit, and the cutting edges were rubbed together for 1 minute to remove any bitterness. The fruit was then peeled, cut into ½-inch pieces, and all seeds were removed. The seed pulp was mixed with another ½ teaspoon of salt and water, and the gelatinous, edible juice was collected. The *C. melo* slices

were combined with the collected juice and left to sit for 30 minutes to allow more juice to be released from the pieces. After this, all the ingredients listed in the table were mixed thoroughly and cooked over medium heat for 10 minutes. Finally, 200 ml of coconut milk was added, and the curry was stirred continuously until a thick gravy formed.

The same recipe was followed for all *Cucumis melo* L. and *Cucumis sativus* L. salad varieties.

Table 3: Ingredients of prepared salad dishes

Ingredients for 800g of <i>Cucumis</i> spp.	Quantity of ingredients
Onion	100 g
Tomato	100 g
Raw Chili	20 g
Coconut milk	50ml
Black pepper	1 tsp
Salt	1tsp

Preparation of *Cucumis* Varieties

Small slices were removed from the top of the drupes, and the cutting edges were rubbed for 1 minute to remove bitterness, following the conventional *Cucumis* cooking method. The drupes were then peeled, washed thoroughly, dried with tissue paper, and cut into ½-inch pieces, with all seeds removed. Onion, tomato, and raw chili were also sliced into small pieces and mixed with the *Cucumis* varieties. Black pepper and salt were added to coconut milk, which was then poured over the salad mixture.

For the *Cucumis melo* fruit varieties, the following preparation was used: the ripe *C. melo* fruit peel was removed, and the interior placental content was discarded. The fleshy part was crushed in a container, mixed with sugar, and consumed as a fruit salad.

Organoleptic Data Analysis

Consumer preferences were evaluated based on five criteria: color, aroma, texture, bitterness, and overall preference. A three-point scoring system was used, where consumers assigned scores of three, two, or one, representing high, medium, and low preference, respectively. The data were analyzed using the FREQ procedure in SAS to explore associations with varietal preferences. For each *Cucumis* variety, a weighted score was calculated by multiplying the raw percentages by the corresponding rank for each taste parameter. Separate weighted scores were assigned for color, aroma, texture, bitterness, and overall preference for each variety.

These weighted scores were then analyzed using Principal Component Analysis (PCA) in Minitab. The resulting PCA biplot, displaying the two main components (PC1 and PC2), along with the Eigenvalue and Scree plots, illustrated how each sensory parameter contributed to the overall variance in consumer preferences across all *Cucumis* varieties.

Optimization of DNA Extraction and PCR

Genomic DNA was isolated from immature *Cucumis* leaves using the CTAB method. The extracted DNA samples were stored at -20°C . PCR was carried out using a Thermal Cycler (TP600, Takara, Otsu Shiga, Japan) with two DNA barcoding markers (*trnH-psbA* and *rbcL*) (S1 Table). The PCR mixture (15 μl) contained $5\times$ Go Taq Green Master Mix (7.5 μl), 10 μM forward and reverse primers (0.5 μl each), and 10 μM spermidine (3.5 μl) (Wan et al., 1993). PCR conditions are provided in S1 Table. The PCR products were purified using the QIAquick® PCR Purification Kit (Catalog No: 28104, Qiagen, Hilden, Germany). The purified PCR amplicons were size-separated using 2.5% agarose gel electrophoresis (Sambrook et al., 2001).

Table 4: Evaluated key plant DNA barcoding markers.

DNA marker	Sequence	Ta (°C)	Reference
<i>trnH-psbA</i>	F- CGCGCATGGTGGATTCAACAATCC	64	Tate and Simpson (2003)
	R- GTTATGCATGAACGTAATGCTC		Sang <i>et al.</i> , (1997)
<i>rbcL</i>	F- ATGTCACCACAAACAGAGACTAAAGC	55	Levin <i>et al.</i> , (2003)
	R- GTAAAATCAAGTCCACCCRCG		Kress and Erickson (2009)

DNA Sequencing

PCR products were purified using the QIAquick PCR Purification Kit (Catalog No: 28104, Qiagen, Hilden, Germany) to ensure clean samples for sequencing. The purified products were sequenced using the Applied Biosystems Genetic Analyzer 3500. The resulting sequence data were then analyzed for phylogenetic comparisons.

Phylogenetic Analysis

Raw sequencing outputs were first visualized in MEGA 7 (Kumar et al., 2016) to assess initial and terminal noise. Separate alignments for each marker (*trnH-psbA*, *rbcL*) were generated using the Clustal W algorithm (Thompson et al., 1994) in MEGA 7. The noisy regions at the beginning and end of the sequences were trimmed, and alignments were manually checked to prevent automated methods from introducing unwanted INDELs. Once the datasets were aligned, they were combined and exported as a single alignment to PAUP 4.0a (Swofford et al., 2009). A UPGMA tree was constructed using uncorrected pairwise distances, with gaps treated proportionally to non-ambiguous SNPs to ensure fair treatment of

INDELs. All substitutions were given equal weight, and the final dendrogram was refined using FigTree v1.4.3 (Rambaut, 2014).

Subsequently, Maximum Likelihood (ML) analysis was performed using RAxML (Stamatakis et al., 2006), with a rapid bootstrap algorithm run for 1000 iterations. The GTRGAMMA evolutionary model (Stamatakis et al., 2008) was selected to best fit the dataset. Bootstrap values were incorporated into the final ML tree, which was further refined using FigTree v1.4.3.

A Bayesian analysis was also conducted to support the ML tree. The best nucleotide substitution model for the dataset was identified using the Akaike Information Criterion (AIC) (Akaike, 1974) and the corrected AIC (AICc) (Cavanaugh, 1997) in the J-model test (Posada, 2008). The selected model parameters were used to build the tree in a Bayesian framework using MrBayes (Huelsenbeck and Ronquist, 2001). Four Markov Chain Monte Carlo (MCMC) simulations were run for 50 million iterations, with one tree sampled every 5000 chains. The first 10% of trees were discarded as burn-in, and a 50% majority-rule consensus tree was constructed with posterior probabilities (PP) calculated for each branch. The congruence between the Bayesian and ML trees was evaluated, and the PP values were added to the nodes of the ML tree. All tree construction and model selection were performed using the CIPRES Science Gateway (Miller et al., 2010).

Divergence Time Calibration

The clade-age calibration method (Matschiner et al., 2017) was used to estimate divergence times. The dataset included two markers (*trnH-psbA* and *rbcL*) from Williams et al. (2017), along with sequences generated in the current study. The BEAUTi program, part of the BEAST v1.10.4 package, was utilized to calibrate the species tree (Drummond and Rambaut, 2007). The two datasets were imported separately and linked through site models and the final tree. To accommodate different evolutionary models for each marker, the BEAST model-averaging algorithm was applied (Bouckaert and Drummond, 2017), and the Gamma Site Model GTR substitution model was used. A strict molecular clock was employed to calibrate divergence times (Drummond et al., 2006). The calibrated Yule model was chosen to model speciation events without considering extinction, which is suitable for phylogenies with recent divergences. The Yule model also integrates fossil data, biogeographic events, and known historical divergence times, providing an absolute time scale for divergence estimates (Heled and Drummond, 2012).

To mitigate potential inaccuracies in divergence time estimations, fossil records were combined with divergence time calibrations from prior studies (Williams et al., 2017). The robustness of the analysis was assessed using the Effective Sample Size (ESS) values in TRACER (Rambaut and Drummond, 2007). Reliable estimates were ensured by verifying that ESS values were above 200 and Autocorrelation Time (ACT) was low. High ESS values and low ACT indicate reliable convergence (Drummond et al., 2006). The final Maximum Clade Credibility (MCC) tree was generated using TreeAnnotator (Drummond and Rambaut, 2007), and further refined in FigTree v1.4.4 (Rambaut, 2014).

Plant Identification and Barcoding

Both *rbcL* and *trnH-psbA* sequences obtained for *Cucumis melo* L. and *Cucumis sativus* L. were used for DNA barcoding and plant identification. The sequences were compared with

known variety sequences available in GenBank. Since species boundaries were unclear for the *Thiyambara* and *SAYON* varieties, both *rbcL* and *trnH-psbA* sequences were used to confirm their species identity, as these varieties were difficult to distinguish morphologically.

To increase the resolution of the phylogenetic trees and improve species discrimination, the *rbcL* and *trnH-psbA* sequences were integrated. This integrated data enhanced the evolutionary insights, providing a more comprehensive overview of the evolutionary relationships and genetic diversity (Chase et al., 2007).

Results

Variation of the Morphological parameters

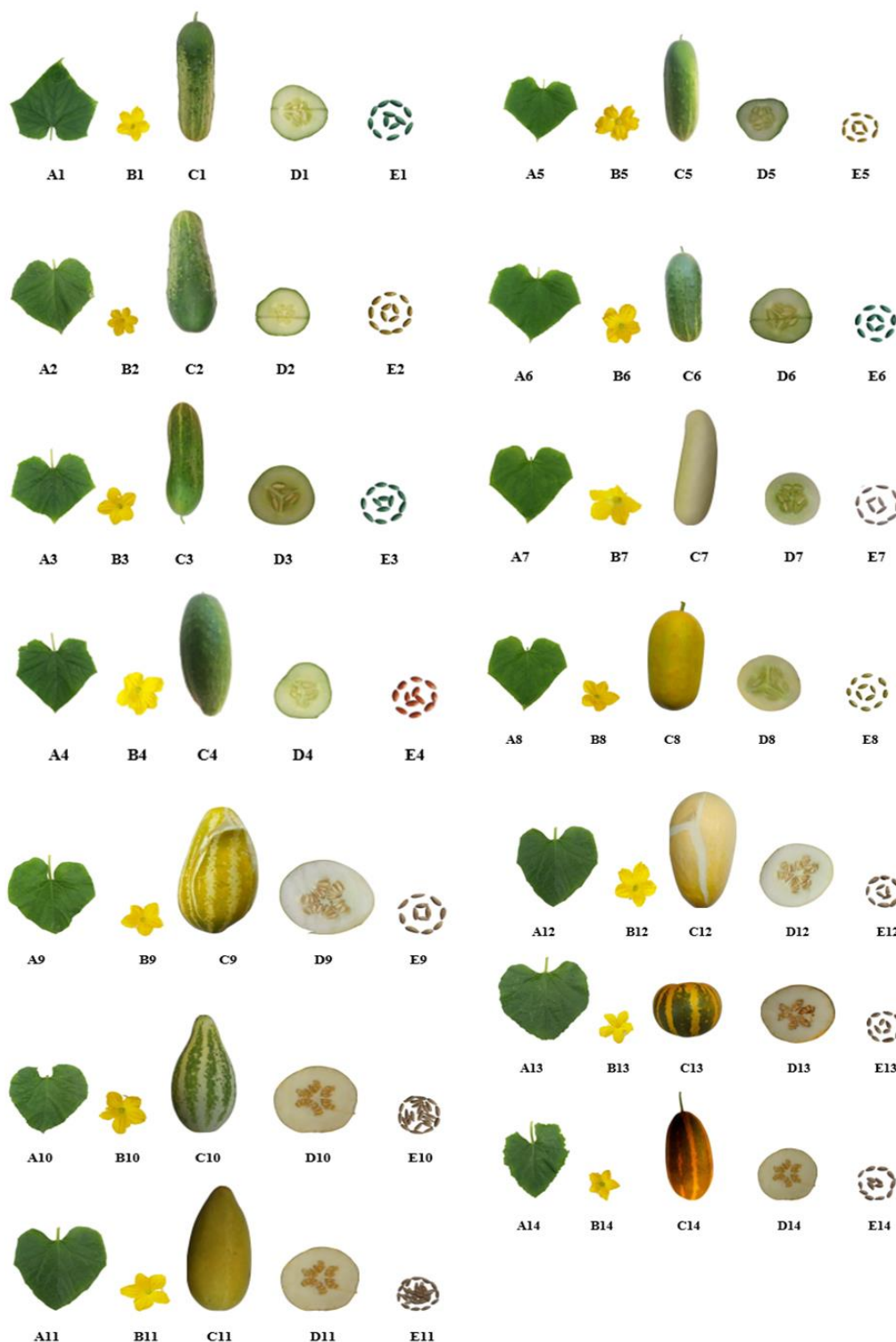


Figure 2: Morphological variation of the reproductive parts (A: leaves; B: flowers; C: fruits appearance; D: cross sections of fruits; E: seeds). The numbers 1–14 indicate Sassy; Wealthy; Fiesty; Ajex; Chandani; Treasure; KW; LY58; GSG; GSS; YSS; WSM; MK; DK.

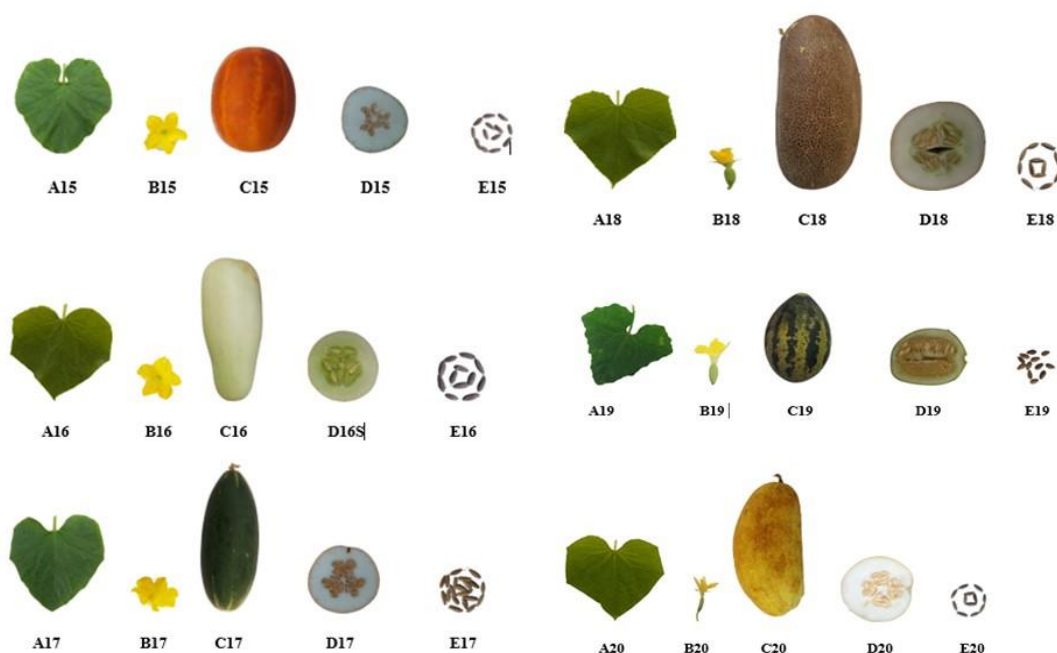


Figure 3: Morphological variation of the reproductive parts (A: leaves; B: flowers; C: fruits appearance; D: cross sections of fruits; E: seeds). The numbers 15–10 indicate SK; DP; LYV; TH; GK; THW.

Growth and Morphological Development

A longitudinal growth analysis (Figure 3.9) assessed the performance variations among different genotypes. Few significant differences in growth rates across the examined varieties were revealed in the longitudinal growth analysis. Varieties *WSM* (V12), *SK* (V15), and *YSS* (V11) exhibited consistently higher mean growth values, respectively 105.45, 104.48, and 102.97. However, these superior growth rates suggest the enhanced adaptability of varieties in the experimental conditions. In contrast, *Sassy* (V1), *MK* (V13), and *DP* (V16) varieties revealed significantly lower mean growth values, respectively 58.26, 79.26, and 82.01. These lower values suggest possible limitations in growth or lack of environmental adaptability within these genotypes.

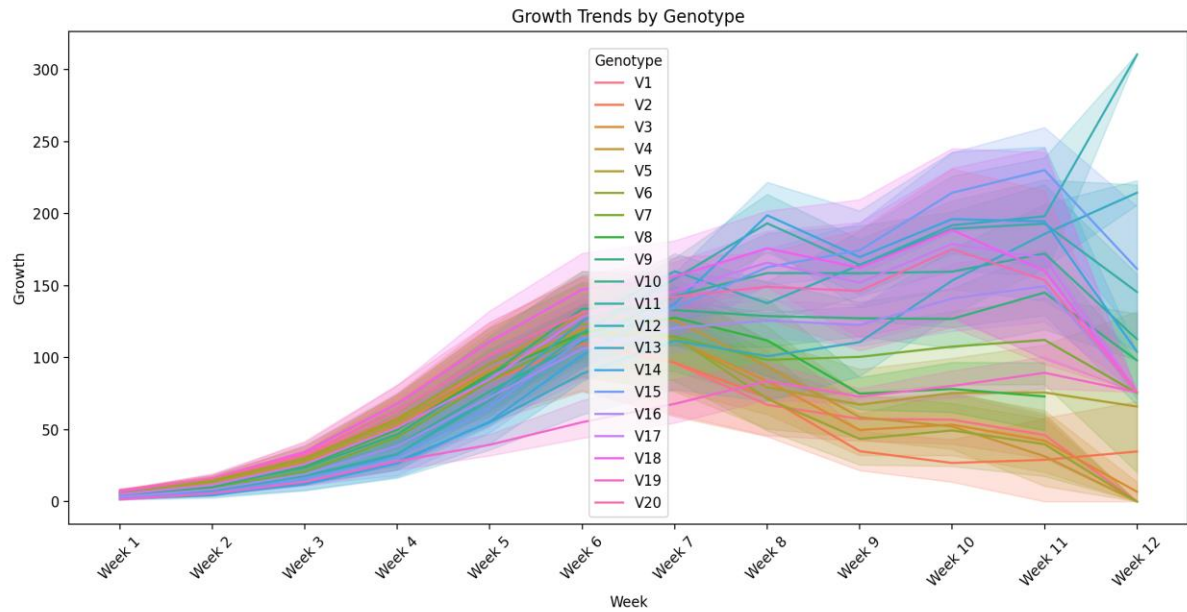


Figure 4: Comparative growth rates of different varieties, showing variability in performance across genotypes. Highlights the superior growth of *WSM* (V12), *SK* (V15), and *YSS* (V11) compared to *Sassy* (V1), *MK* (V13), and *DP* (V16).

In further a significant week-to-week variability in growth patterns across all examined genotypes was revealed in the longitudinal growth analysis. Standard deviations (SD) were applied to quantify the variability. The highest variability (SD = 93.91) was exhibited by the genotype *DK* (V14). Although genotype *DP* (V16) is among the higher-yielding varieties, it showed a comparatively lower variability (SD = 61.63) (Figure 5).

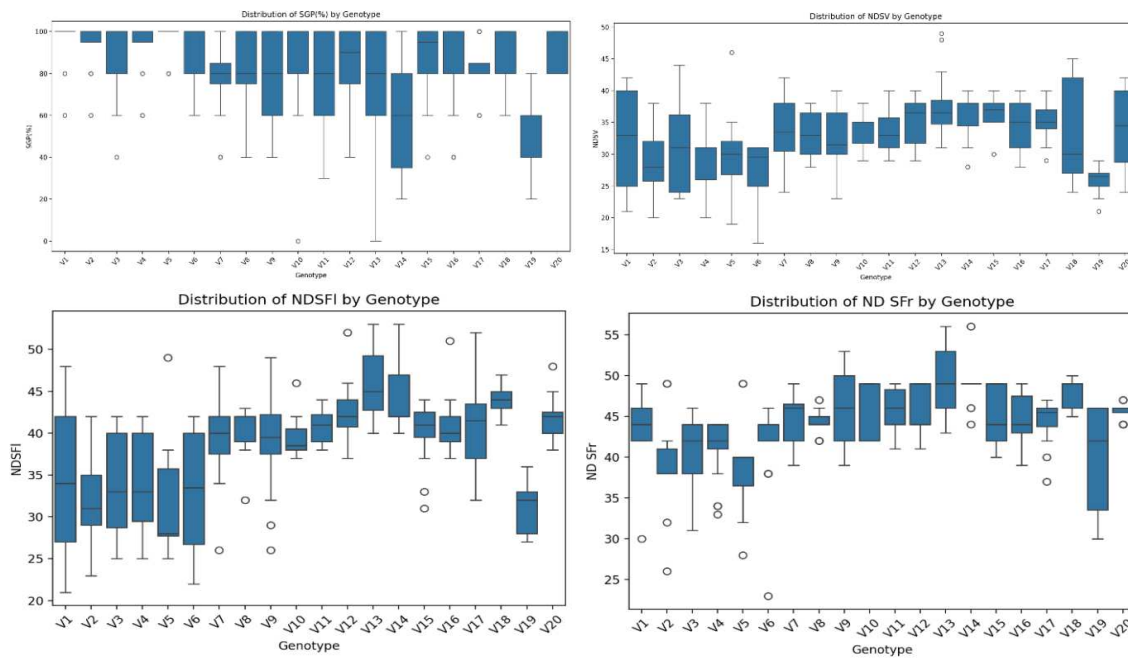


Figure 5: Boxplot analysis of morphological parameters in early developmental stages. Demonstrates the variability in SGP (%), NDSV, NDSFI, and NDSFr

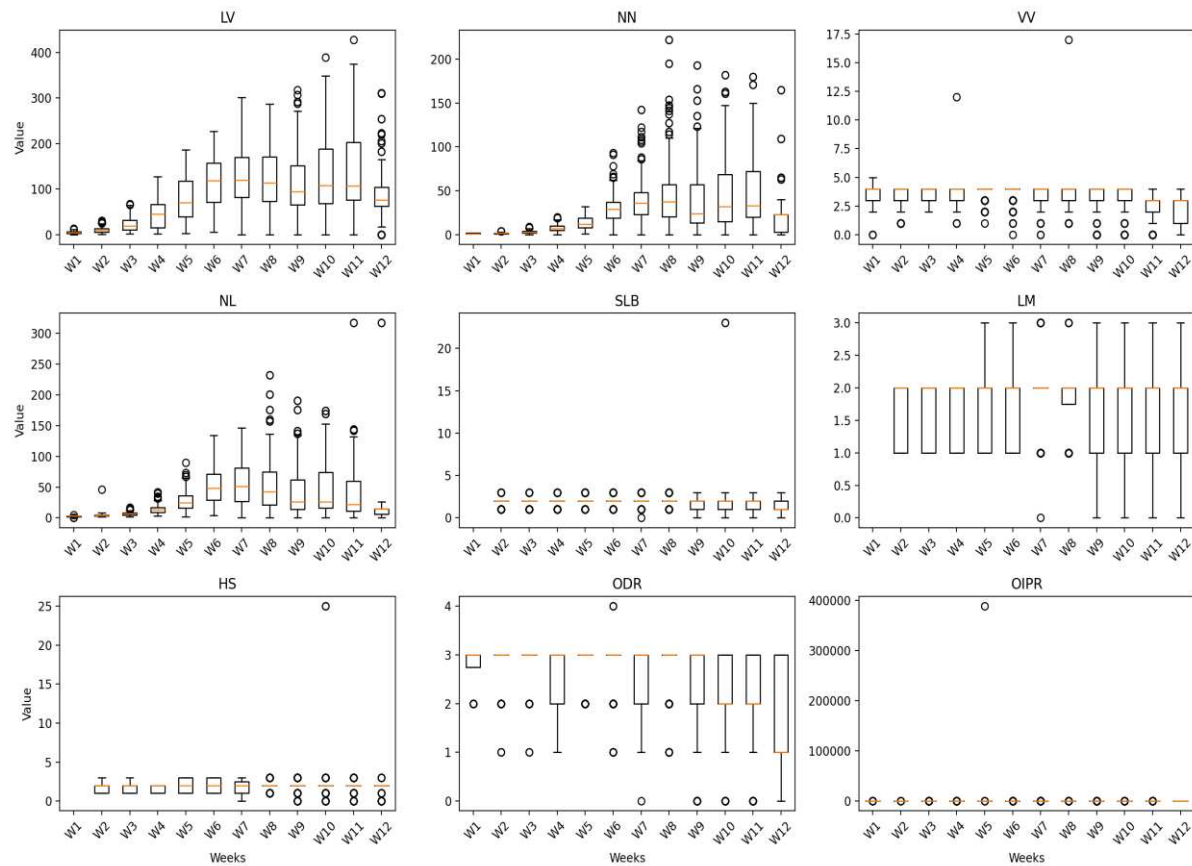


Figure 6 : Boxplot analysis of morphological parameters in early developmental stages. Demonstrates the variability in LV, NN, VV, NL, SLB, LM, HS, ODR, and OPR.

The boxplots (Figure 6) enhanced the understanding of developmental characteristics of the Morphological parameter analysis. The widest distribution was noted in the length of the vine (LV) which was ranging from 0 to 13.2 cm in week 1. Therefore, a considerable variability in early vine development was identified. However, the node number (NN) exhibited a more conservative variation. Nod number was ranging from 1 to 2 nodes in week 1. The consistency in nod number is suggesting a stronger genetic influence for that trait. In the same way, the vigor of the vine (VV) established relatively consistent measurements, clustering around values of 3 to 4, suggesting a stable characteristic across the studied genotypes.

Organoleptic Assessment

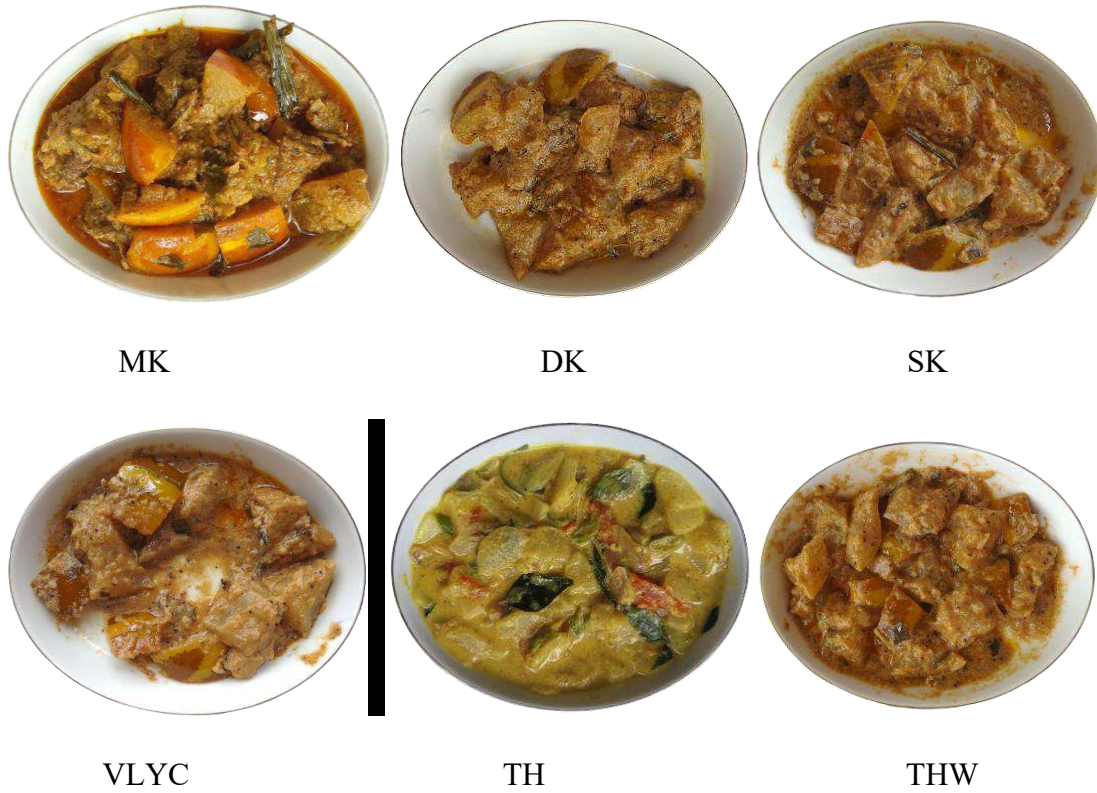


Figure 7: Plate: Culinary preparations subjected to taste panel and subsequent association analyses between sensory parameters and six cooking *Cucumis* varieties.

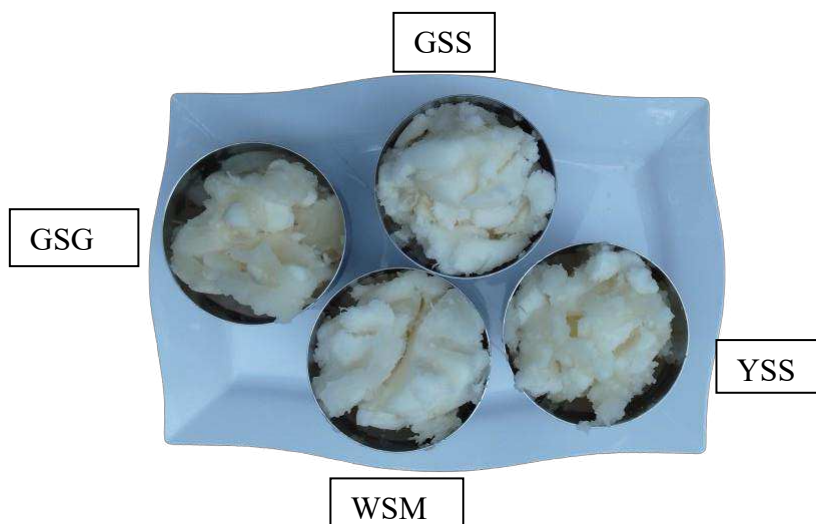


Figure 8: Dish: Fruit salad preparations subjected to taste panel and subsequent association analyses between sensory parameters and *Cucumis melo* L. fruit varieties.

Organoleptic Assessment of Cooking Varieties

Chi-square analysis was used to identify significant relationships between sensory attributes and three culinary preparations. The results showed significant associations between color and overall performance ($\chi^2 = 77.9877$), aroma and overall performance ($\chi^2 = 85.3095$), and texture and overall performance ($\chi^2 = 60.7263$), all with p-values less than 0.05. These findings suggest a strong correlation between these sensory traits, emphasizing the importance of color, aroma, and texture in determining overall sensory preference.

Principal Component Analysis (PCA) was applied to simplify the complex sensory data. It revealed that overall performance (-0.6223), aroma (-0.4908), and color (-0.4766) had the greatest influence on the first principal component (PC1), making them critical factors in the sensory profile. Additionally, a significant correlation was found between texture and bitterness ($\chi^2 = 12.7212$, $p = 0.0127$), highlighting the complex interactions between these attributes. The second principal component (PC2) placed emphasis on bitterness (-0.7673) and texture (0.4921), further underscoring their important relationship (Figure 1).

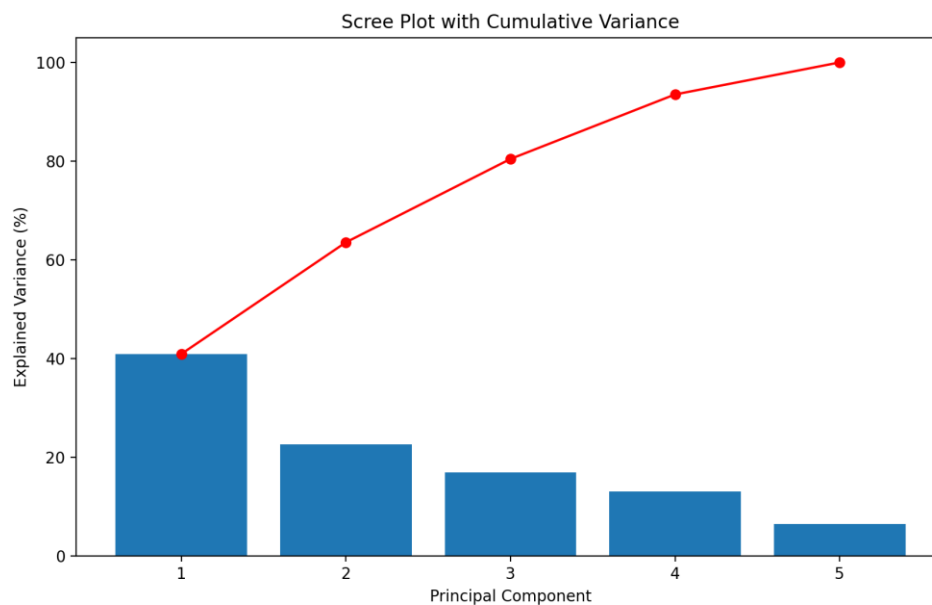


Figure 9: Scree Plot with Cumulative Variance. Scree plot depicting the explained variance for each principal component in the organoleptic assessment analysis. The blue bars represent the percentage of variance explained by individual principal components, while the red line shows the cumulative explained variance. This visualization helps determine the optimal number of principal components to retain for interpreting the sensory attribute data, with components accounting for substantial cumulative variance providing the most meaningful insights into the underlying data structure.

The biplot visualization (Figure 6) provides valuable insights into the relationships between sensory attributes by illustrating their spatial relationships. Attributes that point in similar directions are positively correlated, while those at right angles are independent of one another. This multidimensional approach offers a more comprehensive understanding of the product's sensory profile than analyzing individual attributes alone. A key finding is the strong relationship between overall performance and several sensory attributes, suggesting that consumer satisfaction is influenced by the combined effects of factors such as color,

aroma, texture, and bitterness.

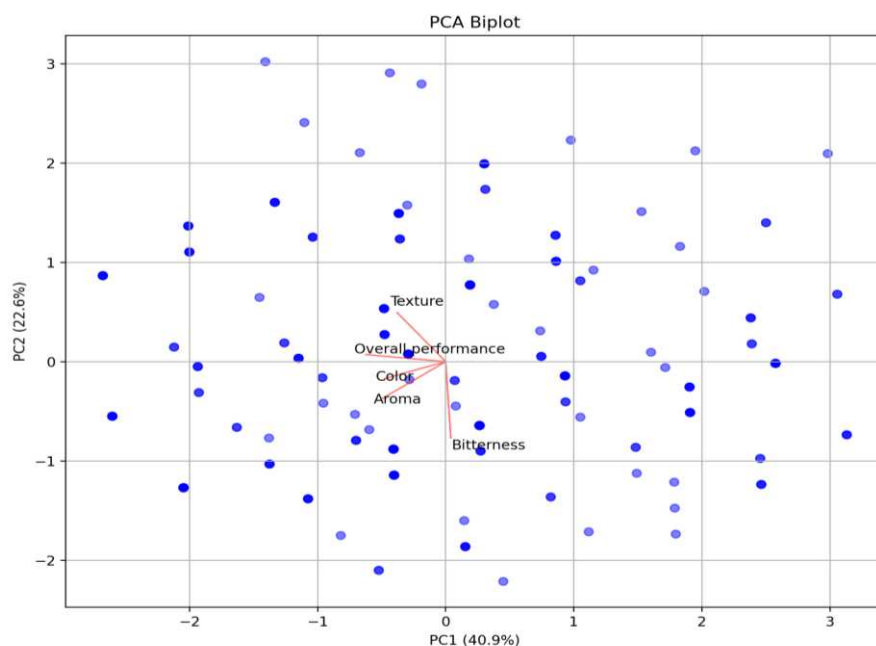


Figure 10: PCA Biplot of Sensory Attributes. Principal Component Analysis (PCA) biplot revealing the relationships between sensory attributes: Color, Aroma, Texture, Bitterness, and Overall Performance. The blue scatter points represent individual data points projected onto the first two principal components (PC1 and PC2). Red arrows indicate the direction and magnitude of each sensory attribute's contribution to the principal components. The length and orientation of these vectors suggest the relative importance and correlation between attributes, with closer-aligned vectors indicating stronger relationships. The axes display the percentage of variance explained by PC1 and PC2, providing context for the dimensionality reduction.

Organoleptic Assessment of Fruit Varieties

Complex sensory characteristics were revealed in Principal Component Analysis (PCA). The first two principal components collectively explained 62.4% of the total variance. Therefore first two principal components delivered a robust representation of the fruit varieties sensory attributes (Figure 10).

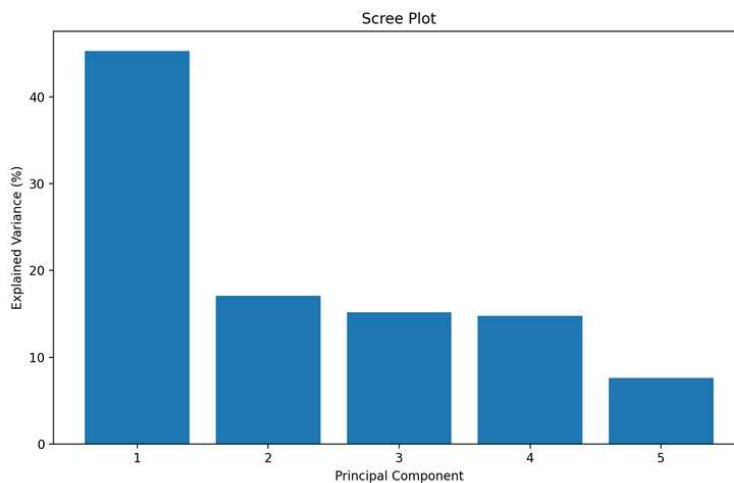


Figure 11: Scree plot illustrating the explained variance for each principal component in the fruit organoleptic assessment. The bar graph displays the percentage of variance explained by individual principal components, providing insight into the dimensionality reduction process. This visualization helps researchers identify the most significant components that capture the underlying variability in sensory attributes, crucial for understanding the data's structure and complexity.

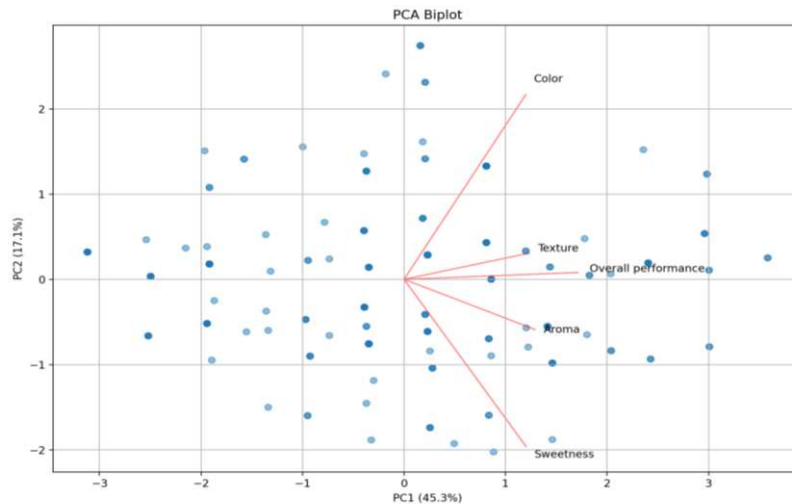


Figure 12: PCA Biplot. Principal Component Analysis (PCA) biplot revealing the multivariate relationships between key sensory attributes: Color, Aroma, Texture, Sweetness, and Overall Performance. The scatter plot represents individual data points projected onto the first two principal components (PC1 and PC2). Red arrows indicate the direction and magnitude of each attribute's contribution to the principal components, with arrow length and orientation suggesting the relative importance and potential correlations between different sensory characteristics.

A strong correlation between overall performance and sweetness was identified in biplot analysis (Figure 3.65). Accordingly, these two attributes were closely associated in the sensory evaluation. Sweetness is one of key parameters that determine the fruit quality. Therefore, sweetness showed a positive impact on overall performance. Consequently, sweetness should be considered as the most elite trait in breeding programs and fruit selection. In further, Color and texture showed weak correlations with other sensory attributes. Therefore, they can be adjusted separately without greatly affecting the overall sensory profile.

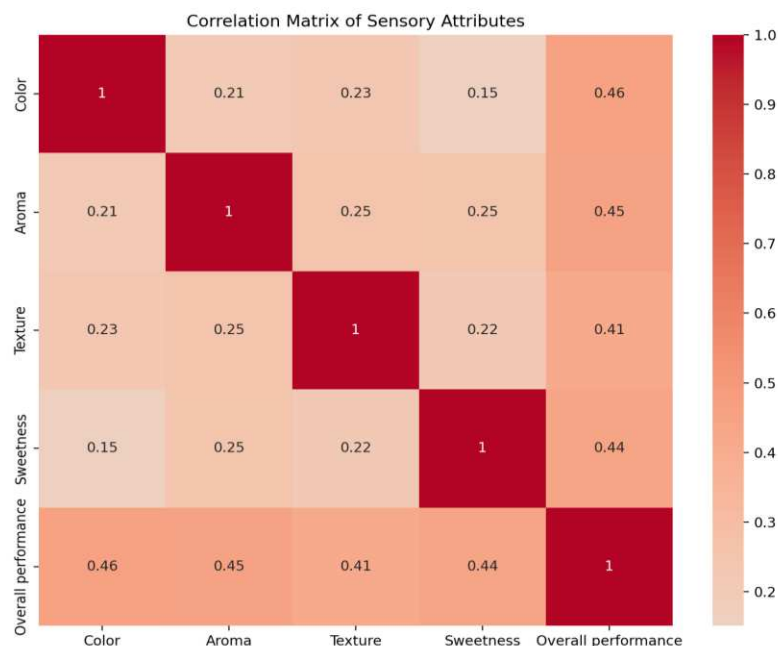


Figure 13 :Correlation Matrix Heatmap. Correlation matrix heatmap displaying the inter-relationships between sensory attributes using a color-coded scale. The visualization uses a diverging color palette (cool to warm) to represent the strength and direction of correlations between Color, Aroma, Texture, Sweetness, and Overall Performance. Numerical annotations provide precise correlation coefficients, allowing for a detailed interpretation of how different sensory attributes relate to one another in the fruit organoleptic assessment.

In advance, the correlation matrix (Figure 10) also supported to the PCA results. According to Correlation Matrix Heatmap (Figure 10), strong correlation between overall performance and sweetness was confirmed. Even though sweetness is considered as the most preferred sensory attribute, other attributes also added unique features to the fruit's sensory experience

Organoleptic Assessment of Salad Varieties

Principal Component Analysis (PCA) reveals distinct sensory profiles for different salad types. The scatter plot demonstrates a clear separation of samples along the principal components, highlighting noticeable differences in sensory characteristics. The first two principal components capture the majority of the variation, offering a detailed understanding of consumer preferences.

The correlation matrix further illustrates how sensory attributes influence overall performance. Color shows a moderate positive correlation (0.326), emphasizing the importance of visual appeal in consumer perception. Bitterness exhibits a strong negative correlation (-0.493), indicating that higher bitterness significantly diminishes consumer preference. Aroma and texture have weaker positive correlations (0.235 and 0.202, respectively), subtly contributing to the overall sensory experience (Figures 11 & 12).

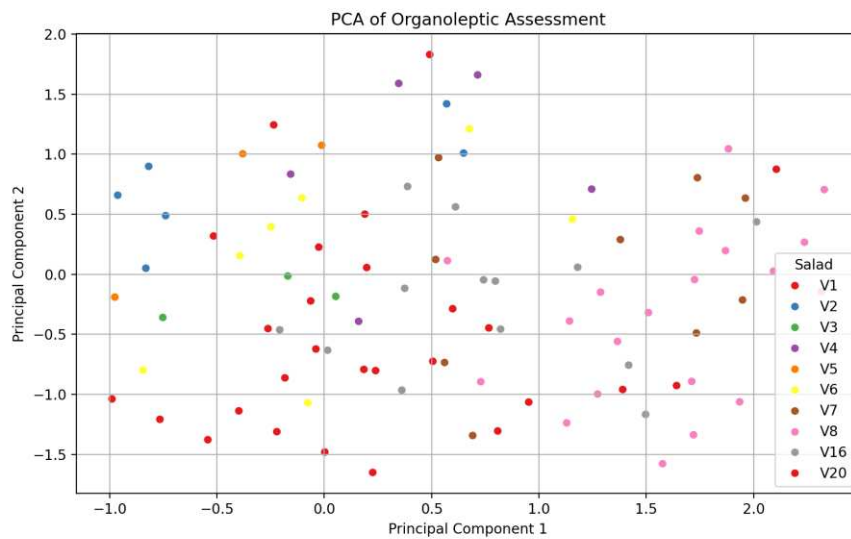


Figure 14: PCA Scatter Plot. Principal Component Analysis (PCA) scatter plot revealing the distribution of different salad variants across the first two principal components. Each point represents a salad sample, color-coded by salad type to illustrate the sensory attribute variations. The plot demonstrates the multivariate differences between salad samples, showing how they cluster or diverge based on their organoleptic characteristics.

The scree plot indicates that the first two principal components account for a significant portion of the total sensory variation. The sharp decline in explained variance after the second component suggests that these two components effectively capture the primary sensory differences in the salad samples, providing a strong foundation for understanding consumer preferences.

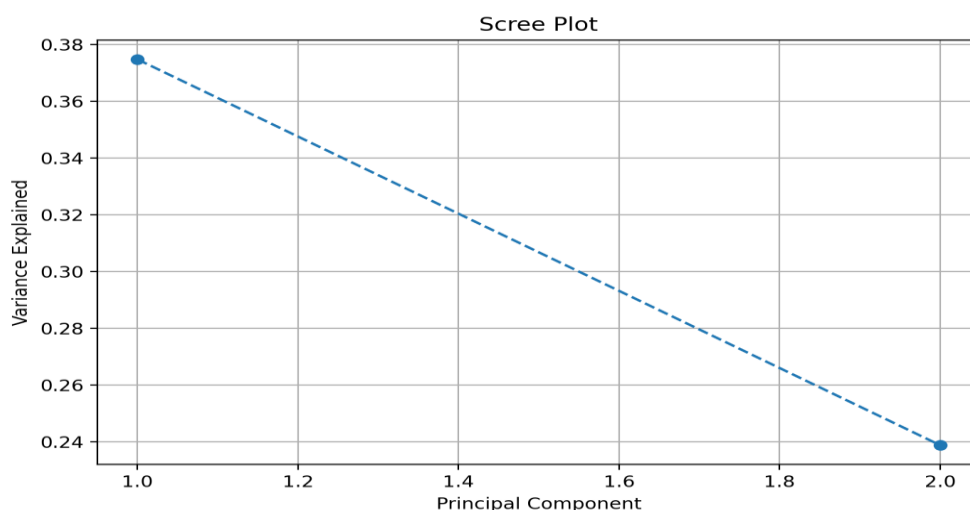


Figure 15: Scree Plot. Scree plot illustrating the proportion of variance explained by each principal component in the salad organoleptic assessment. The line graph displays the cumulative explained variance, helping identify the most significant components that capture the underlying variability in sensory attributes.

The correlation biplot (Figure 13) illustrates how sensory attributes interact and influence one another. The lengths and directions of the vectors reflect the importance of each attribute, while attributes positioned at near-right angles are largely independent. This suggests that each sensory characteristic contributes uniquely to the overall perception of the salad.

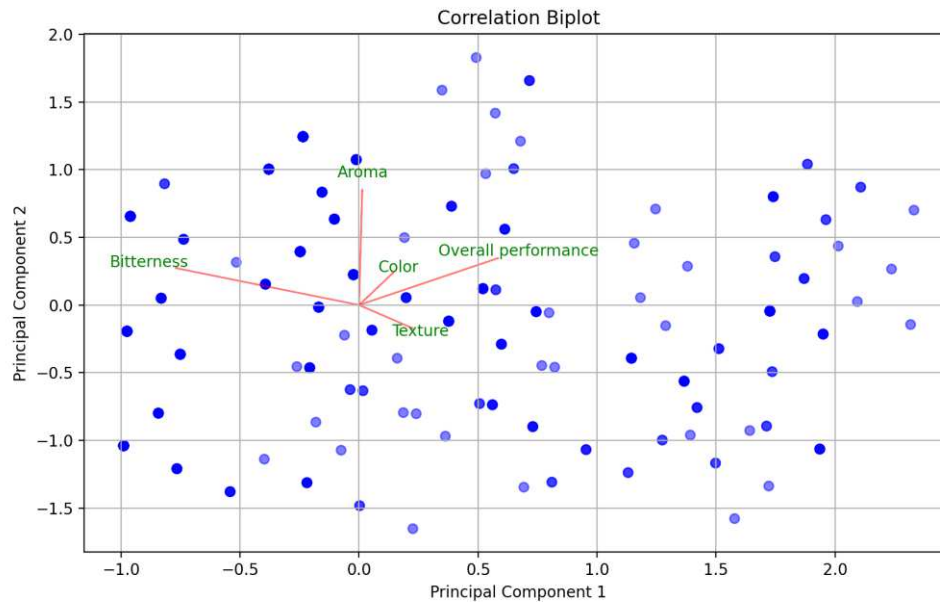


Figure 16: Correlation Biplot. Correlation biplot depicting the relationships between sensory attributes and the first two principal components. Blue scattered points represent individual salad samples, while red arrows indicate the direction and magnitude of each sensory attribute's contribution to the principal components. The text labels show the attributes' positions, revealing how different sensory characteristics are interrelated and contribute to the overall variance in the dataset.

RESULTS AND DISCUSSION

DNA extraction and PCR results

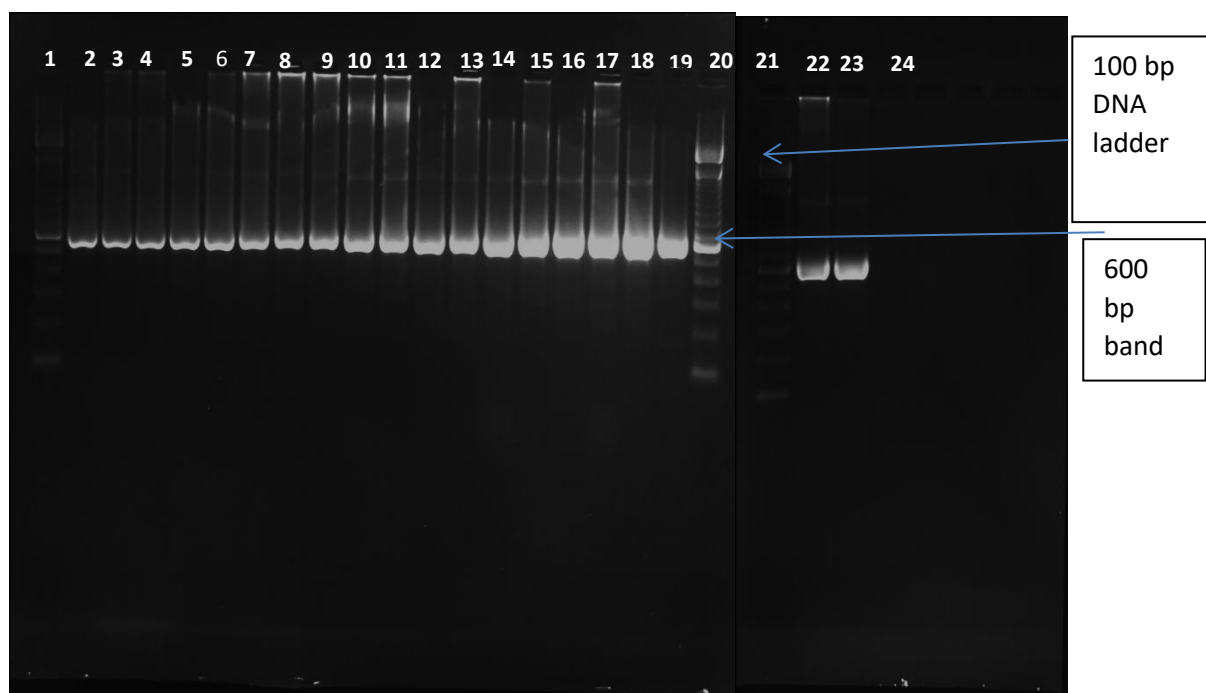


Figure 17: Visualization of the *rbcL* PCR products for *Cucumis* varieties DNA, obtained from fresh leave samples. 1: molecular marker (100 bp ladder), 2: Sassy, 3: Wealthy, 4: Fiesty, 5: Ajex, 6: Chandani, 7: Treasure, 8: KW, 9: LY58, 10: GSG, 11: GSS, 12: YSS, 13: WSM, 14: MK, 15: DK, 16: SK, 17: DP, 18: LYV, 19: TH, 20: molecular marker (100 bp ladder), 21: molecular marker (100 bp ladder), 22: GK, 23: THW, 24: Distilled water.

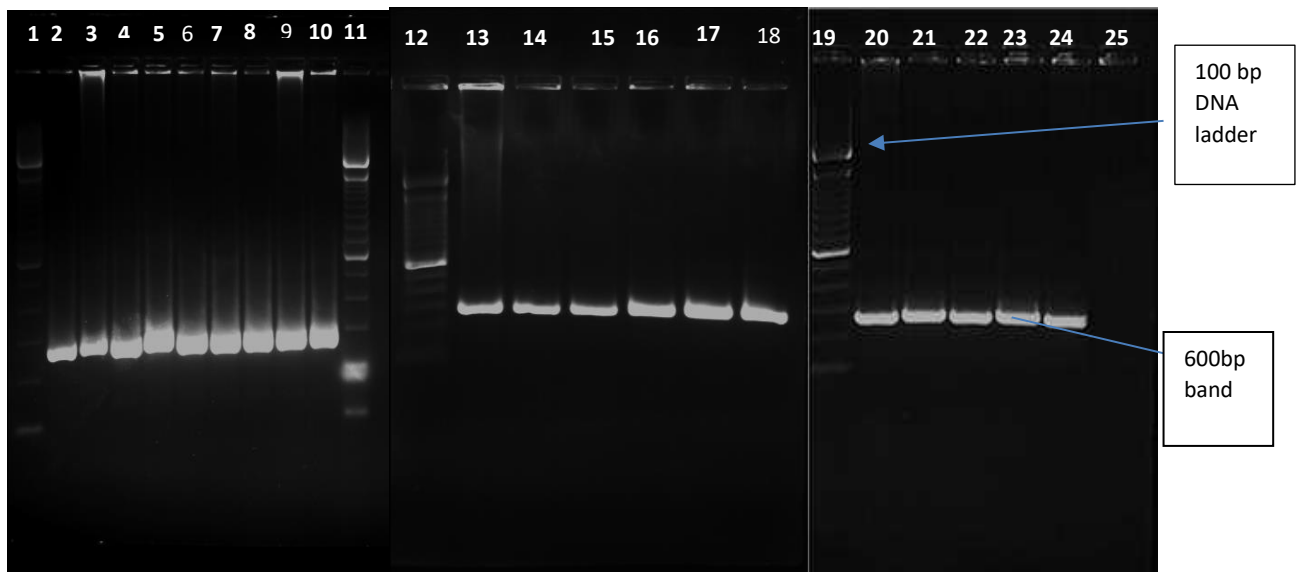
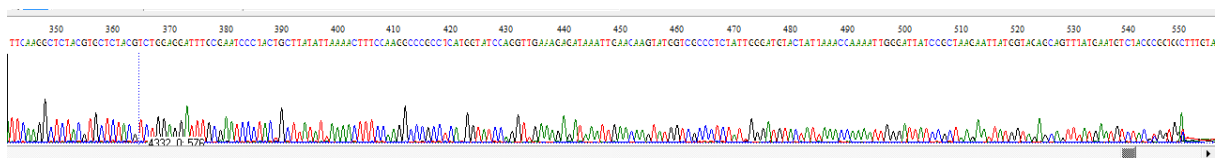


Figure 18: Visualization of the *psbA-trnH* PCR products for *Cucumis* varieties DNA, obtained from fresh leave samples. 1: molecular marker (100 bp ladder), 2: Sassy, 3: Wealthy, 4: Fiesty, 5: Ajex, 6: Chandani, 7: Treasure, 8: KW, 9: LY58, 10: GSG, 11: molecular marker (100 bp ladder), 12: molecular marker (100 bp ladder), 13: GSS, 14: YSS, 15: WSM, 16: MK, 17: DK, 18: SK, 19: molecular marker (100 bp ladder), 20: DP, 21: LYV, 22: TH, 23: GK, 24: THW, 25: Distilled water PCR.

Both gel images for the *rbcL* and *psbA-trnH* PCR products show distinct DNA bands, with minimal smearing, indicating clean and specific results. The 100 bp ladder serves as a molecular ruler, helping to determine the sizes of the DNA fragments in the *psbA-trnH* and *rbcL* PCR products. The DNA bands obtained from the *rbcL* PCR products are approximately 600 bp, as indicated by the 600 bp marker on the ladder. In contrast, the DNA bands from the *psbA-trnH* PCR products are smaller than the 600 bp marker, suggesting that the *psbA-trnH* bands are less than 600 base pairs in length.

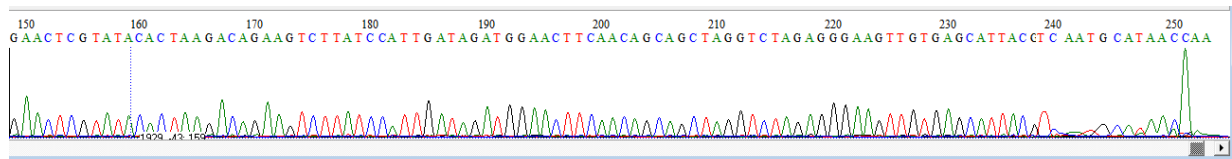
DNA sequencing results

(A)



(A): a part of the chromatogram obtained in sequencing *rbcL* PCR product of VI - Sassy variety (a long readable chromatogram with the DNA sequence).

(B)



(B): a part of the chromatogram obtained in sequencing *trnH-psbA* PCR product of *VI - Sassy* variety (a relatively short (260 bp) readable chromatogram with the DNA sequence).

Figure 19: DNA sequencing of *rbcL* and *trnH-psbA* pop sets of *Cucumis melo* L. and *Cucumis sativus* L.

DNA sequencing results for the *rbcL* and *trnH-psbA* regions were providing readable chromatograms for both loci, though the *psbA-trnH* sequence was shorter (260 bp). However gained *rbcL* sequence was longer (560 bp).

Alignment of DNA sequences

Some part of the sequence alignment results received from MEGA 11 for both *rbcL* and *trnH-psbA* regions are given bellow. The same alignment results were used for generating the Phylogenetic trees and assess evolutionary divergence.



Figure 20: DNA sequence alignment diagram for *rbcL* region of the chloroplast genome of *Cucumis melo* L. and *Cucumis sativus* L. cultivars. The alignment was obtained by using MEGA 11 software. The DNA sequence differences shown in this illustration were used to draw the genetic dissimilarity diagram (i.e. dendrogram) shown in Diagram 1.

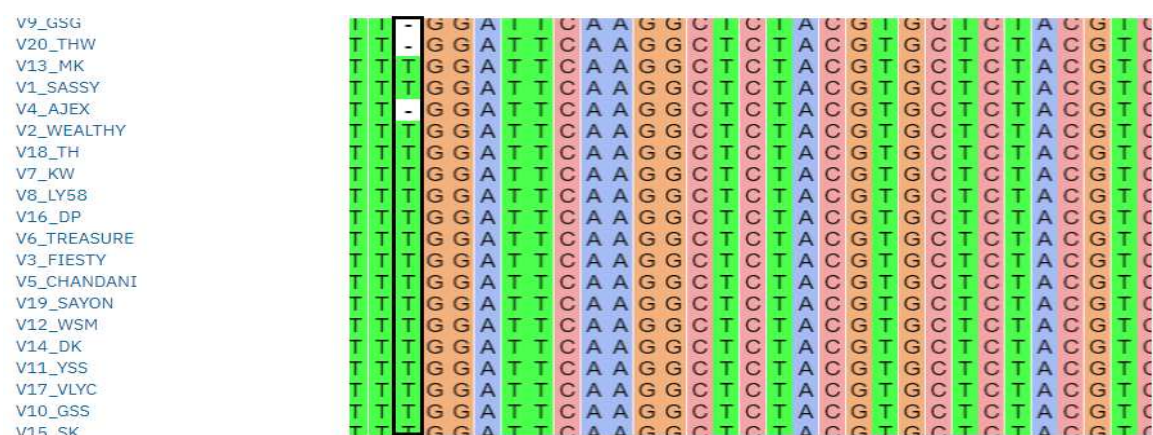


Figure 21: DNA sequence alignment diagram for *trnH-psbA* region of the chloroplast genome of *Cucumis melo* L. and *Cucumis sativus* L. cultivars. The alignment was obtained by using MEGA 11 software. The DNA sequence differences shown in this illustration were used to draw the genetic dissimilarity diagram (i.e. dendrogram) shown in Diagram 2.

The genetic diversity among *Cucumis* varieties in Sri Lanka

The constructed phylogenetic tree, based on the *rbcL* and *trnH-psbA* regions of Sri Lankan *Cucumis* species, highlights the evolutionary relationships among the *Cucumis* germplasm in Sri Lanka. Each branch of the tree represents a distinct lineage, while the nodes, where branches converge, indicate common ancestors. The numerical values near the branches are bootstrap values, which provide statistical support for the groupings (or clades). These values are essential for assessing the reliability of each node in the tree.

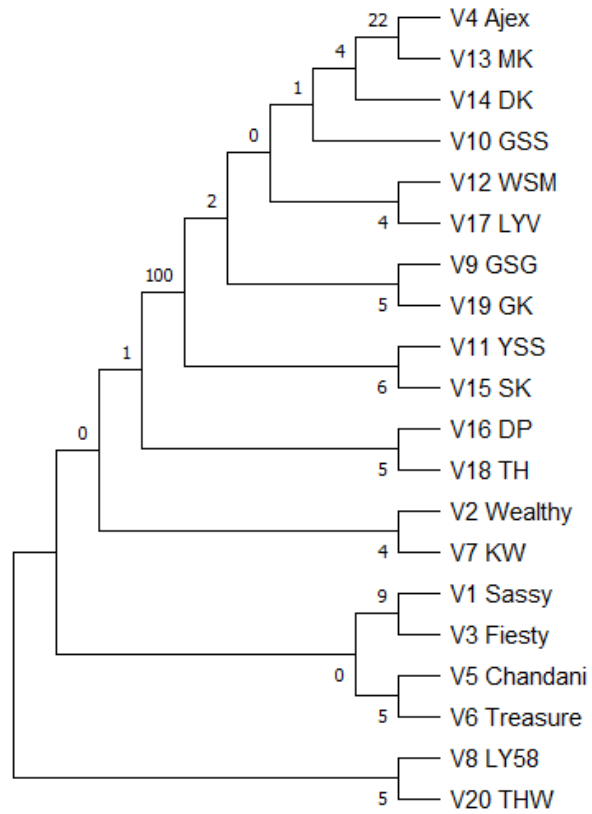


Figure 22: The genetic dissimilarity diagram (i.e. dendrogram) for *rbcL* region of the chloroplast genome of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* cultivars L.

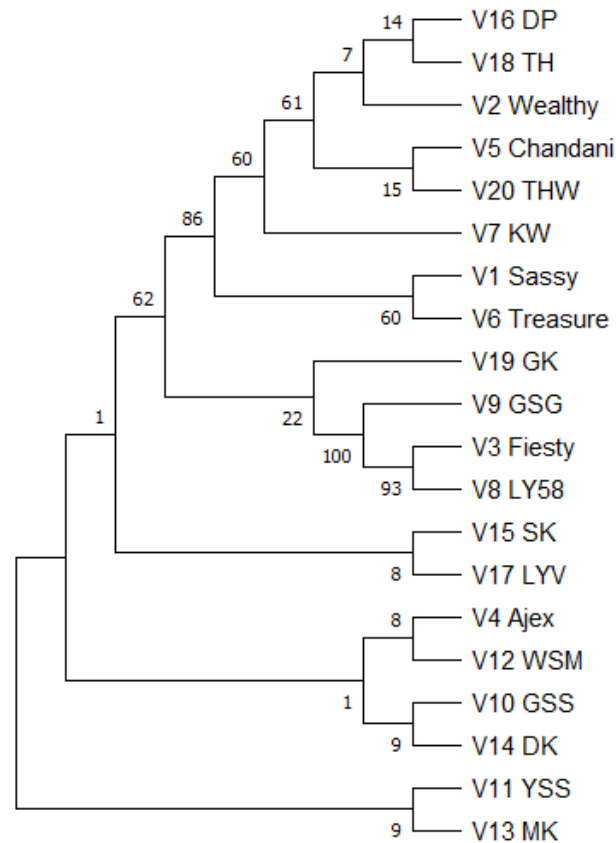


Figure 23: The genetic dissimilarity diagram (i.e. dendrogram) for *trnH-psbA* region of the chloroplast genome of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* cultivars L.

DNA Marker Polymorphism and Genetic Diversity

Informativeness of the DNA Markers

ClustalOmega alignment results revealed that the majority of the aligned DNA sequences did not show gaps or mismatches. However, significant variable regions were identified, indicating areas with higher genetic diversity. A few gaps and mismatches were observed, highlighting regions with lower evolutionary constraints. These sequences may experience mutations without causing significant biological consequences.

When two sequences were closely aligned, it suggested that they shared a more recent common ancestor. For example, the *Sassy* / *Fiesty* and *DK* / *THW* sequence pairs showed close alignment, indicating a close evolutionary relationship between these pairs.

A substantial number of single nucleotide polymorphisms (SNPs) were identified in the alignment results, which were highlighted in the following diagrams. Based on the SNP/INDEL profile, we determined that the combined *rbcl* and *trnH-psbA* markers are the most reliable and informative DNA barcodes. These markers exhibit high inter-specific

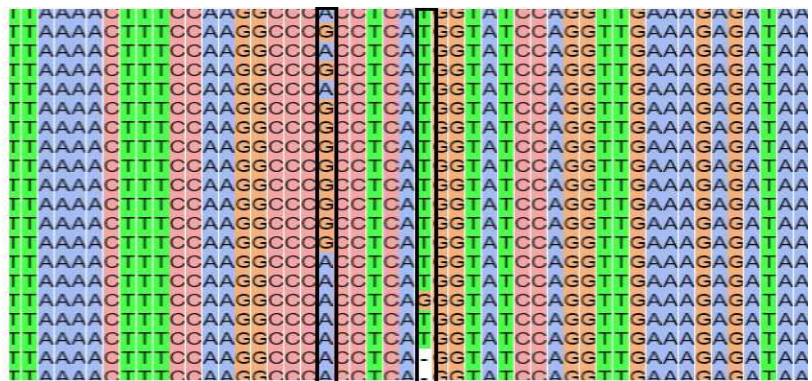
sequence variation and provide clear, unambiguous results. The high efficacy of *rbcL* in species identification and delimitation is also supported by previous findings (Kress and Erickson, 2007).

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK



Base location 394

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK



Base location 420,426

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK



Base location 548

V9_GSG	C	A	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	G	G	A	T	T	C	A	A	A	G	C	
V20_THW	C	C	A	A	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	G	G	A	T	T	C	A	A	A	G	C	
V13_MK	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
V1_SASSY	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V4_AJEX	A	A	C	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V2_WEALTHY	A	C	A	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V18_TH	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V7_KW	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V8_LY58	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V16_DP	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V6_TREASURE	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V3_FIESTY	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V5_CHANDANI	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V19_SAYON	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V12_WSM	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V14_DK	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V11_YSS	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V17_VLYC	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V10_GSS	A	G	A	G	A	C	T	A	A	G	C	A	A	G	T	G	T	T	T	T	G	G	A	T	T	C	A	A	A	G	C
V15_SK	A	C	A	A	C	T	A	A	A	C	C	A	A	A	C	T	C	T	T	T	G	C	A	T	T	C	A	A	A	C	

Base locations 660 to 680

V9_GSG	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V20_THW	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V13_MK	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V1_SASSY	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V4_AJEX	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V2_WEALTHY	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V18_TH	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V7_KW	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V8_LY58	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V16_DP	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V6_TREASURE	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V3_FIESTY	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V5_CHANDANI	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V19_SAYON	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V12_WSM	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V14_DK	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V11_YSS	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V17_VLYC	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V10_GSS	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G
V15_SK	G	G	G	G	C	C	A	G	C	T	G	T	A	G	C	T	G	C	C	T	G	G	A	A	T	C	C	T	T	C	A	C	T	G

Base locations 811,812,815

V9_GSG	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V20_THW	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V13_MK	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V1_SASSY	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V4_AJEX	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V2_WEALTHY	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V18_TH	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V7_KW	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V8_LY58	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V16_DP	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V6_TREASURE	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V3_FIESTY	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V5_CHANDANI	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V19_SAYON	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V12_WSM	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V14_DK	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V11_YSS	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V17_VLYC	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V10_GSS	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T
V15_SK	T	A	T	T	G	C	T	T	A	T	G	T	A	G	C	T	T	A	T	C	C	C	C	T	C	C	C	T	A	G	A	C	C	T

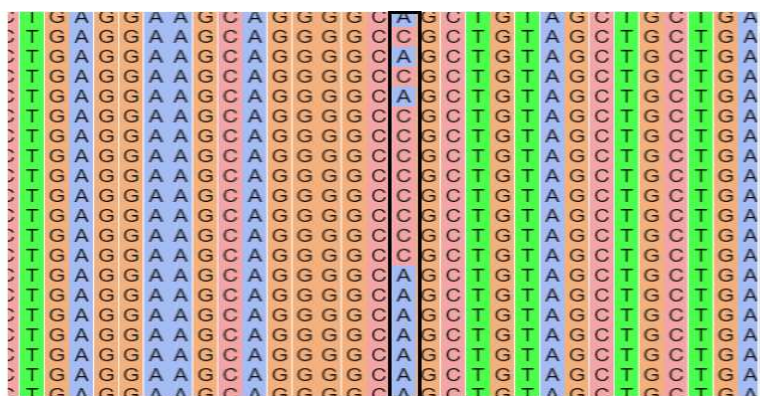
Base location 959

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SKS

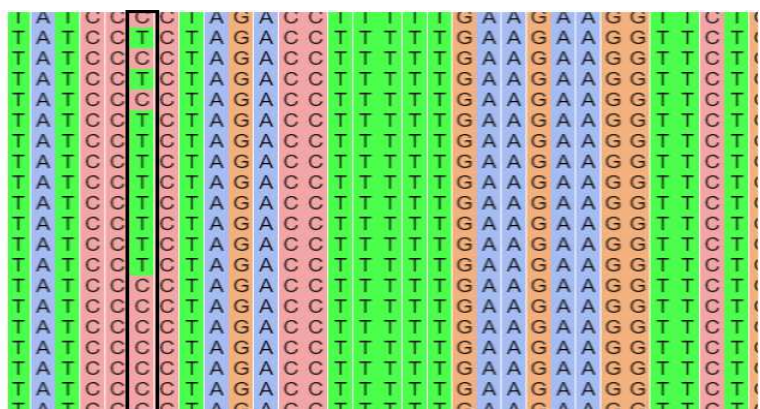
Page 33 of 51

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK



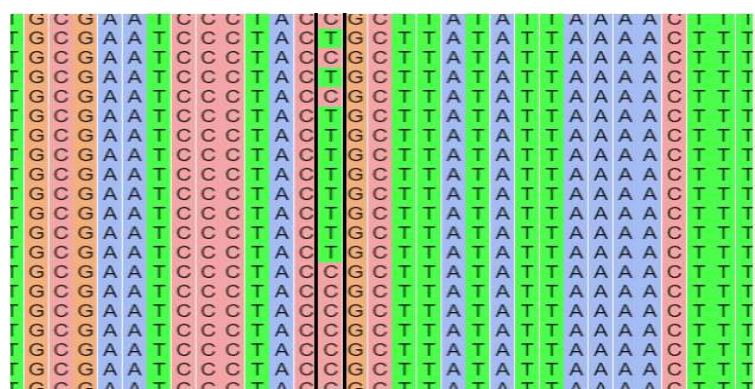
Base location 1343

V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK

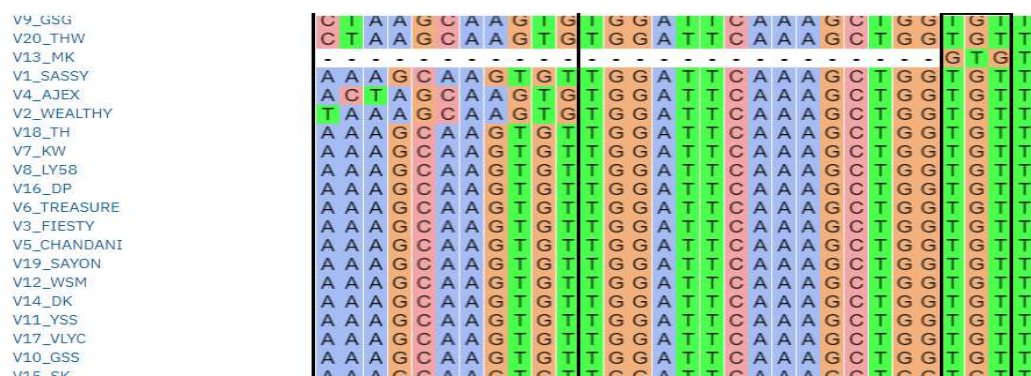


Base location 1489

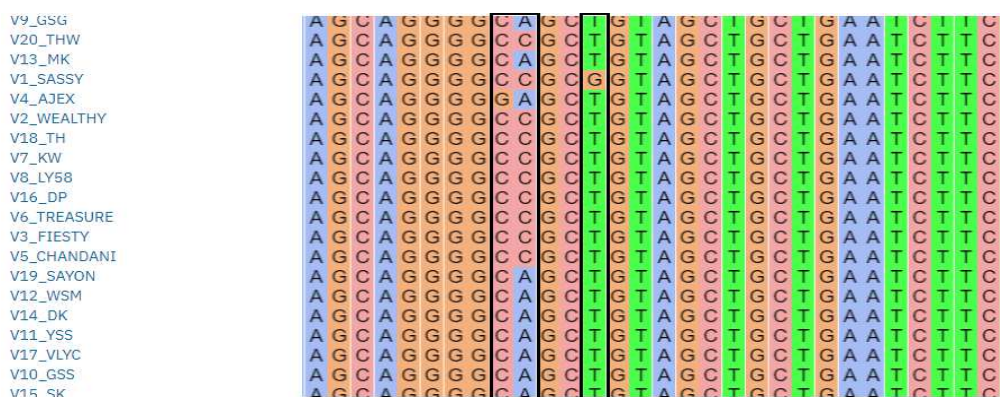
V9_GSG
V20_THW
V13_MK
V1_SASSY
V4_AJEX
V2_WEALTHY
V18_TH
V7_KW
V8_LY58
V16_DP
V6_TREASURE
V3_FIESTY
V5_CHANDANI
V19_SAYON
V12_WSM
V14_DK
V11_YSS
V17_VLYC
V10_GSS
V15_SK



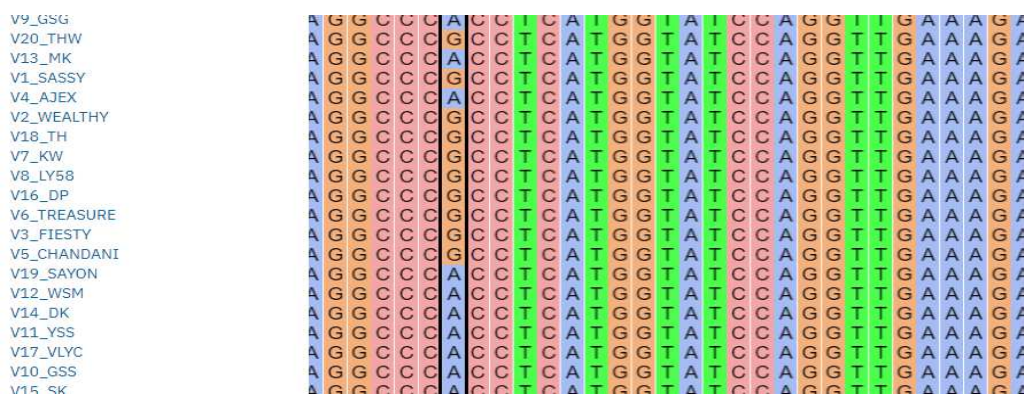
Base location 1603



Base locations 1878 to 1908



Base locations 2021 and 2024



Base location 2309

Figure 24: Base locations of SNPs, INDELs, and shared bases are highlighted for easy visualization for *rbcL* and *trnH-psbA* locus in *Cucumis melo* L. and *Cucumis sativus* germplasms in Sri Lanka.

The alignment analysis identified both conserved and variable regions, with observed single nucleotide polymorphisms (SNPs) indicating inter-specific sequence variation. The high efficacy of the *rbcL* and *trnH-psbA* loci in species identification aligned with findings from

Kress and Erickson (2007), confirming the reliability of these regions as DNA barcodes for species delimitation in *Cucumis*.

Plant Identification and Barcoding

Phylogenetic Analysis

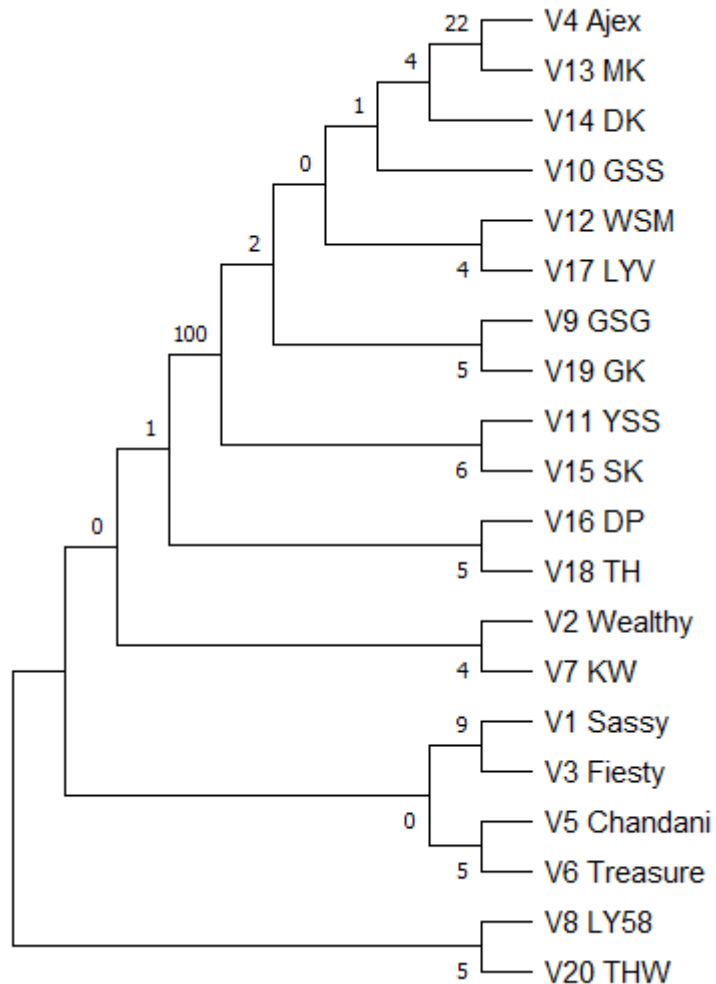


Figure 25: Molecular Phylogenetic affinity analysis of *rbcL* gene sequences of *Cucumis melo* L and *Cucumis sativus* L. plants of Sri Lanka. The evolutionary history inferred using the Neighbor-Joining method. The tree drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances computed using the Maximum composite likelihood method and are in the units of the number of base substitutions per site. The analysis included 20 nucleotide sequences of 20 plants and evolutionary analyses were conducted in MEGA11.

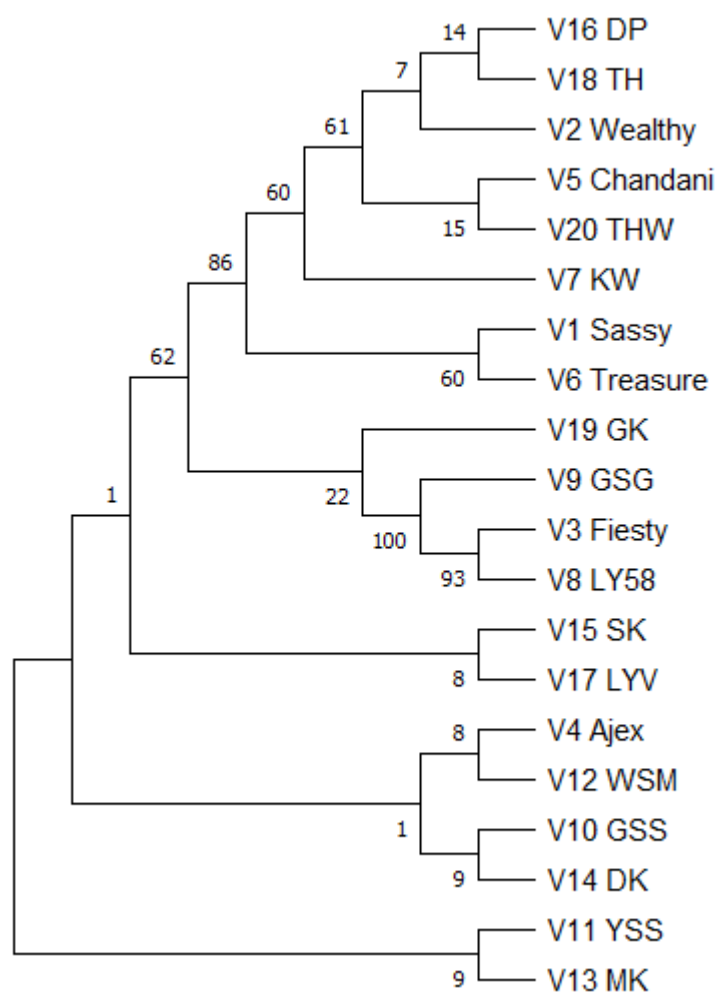


Figure 26: Molecular Phylogenetic affinity analysis of *trnH-psbA* gene sequences of *Cucumis melo* L. and *Cucumis sativus* L. plants of Sri Lanka. The evolutionary history inferred using the Neighbor-Joining method. The tree drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances computed using the Maximum composite likelihood method and are in the units of the number of base substitutions per site. The analysis included 20 nucleotide sequences of 20 plants and evolutionary analyses were conducted in MEGA11.

DNA Marker Polymorphism and Genetic Diversity

Phylogenetic Analysis Using *rbcL* and *trnH-psbA*

Two major clades were identified in the *rbcL* phylogenetic tree (Figure 23). A strong genetic similarity was observed among the varieties *GSG*, *GK*, *YSS*, *SK*, *DP*, and *TH*, which clustered together in a well-supported group (bootstrap = 100). In contrast, another major clade consisted of varieties including *Ajax*, *MK*, *DK*, *GSS*, *WSM*, and *LYV*, which formed several subclusters with moderate to low bootstrap support (0–22). Varieties such as *Sassy*, *Fiesty*, *Chandani*, *Treasure*, *LY58*, and *THW* grouped together but showed relatively lower bootstrap values.

The *trnH-psbA* tree revealed a more resolved structure with higher bootstrap support across major branches. The topology of the trees (Figure 22, Figure 23) differed in several places. For instance, a strongly supported cluster grouping *Chandani*, *Treasure*, *LY58*, *THW*, and *Fiesty* in the *rbcL* tree was altered in the *trnH-psbA* tree. The clade containing *GSG*, *GK*, *YSS*, *SK*, *DP*, and *TH* showed minor topological differences between the two trees. Genotypes such as *Sassy*, *Wealthy*, and *KW* exhibited variability in their positioning across the two trees, possibly due to differences in the resolution power of the two markers.

The *trnH-psbA* tree provided better resolution and discriminatory power among genotypes, offering higher bootstrap support across more nodes. In contrast, the *rbcL* gene showed limited divergence across closely related *Cucumis* accessions, likely due to its more conserved nature. The consistent clustering of varieties such as *GSG*, *GK*, *YSS*, *SK*, *DP*, and *TH* suggests these may represent a common ancestral lineage or share similar geographic/ecological backgrounds.

Plant Identification, Barcoding, and Phylogenetic Analysis Using Integrated Marker Systems (Concatenated Method)

The concatenated phylogenetic tree based on the combined *rbcL* and *trnH-psbA* markers (Figure 24) exhibited increased bootstrap values at most internal nodes, indicating stronger support for the tree's structure. Varieties such as *GSS*, *LYV*, *WSM*, *DK*, and *SK* consistently clustered together with strong bootstrap support (89, 32, and 52, respectively). The cluster of *Ajex*, *GSG*, and *YSS* was retained with better resolution (bootstrap = 87). Varieties including *Treasure*, *Fiesty*, and *Chandani* formed a more stable clade with a bootstrap value of 87.

Clusters such as *GSS*, *LYV*, *WSM*, *DK*, *SK*, and *Ajex*, *GSG*, *YSS* appeared robust across all datasets, supporting their shared ancestry or genetic closeness. Overall, the concatenated phylogenetic analysis of *rbcL* and *trnH-psbA* DNA barcodes enhanced the robustness and resolution of genetic relationships among *Cucumis* accessions in Sri Lanka.

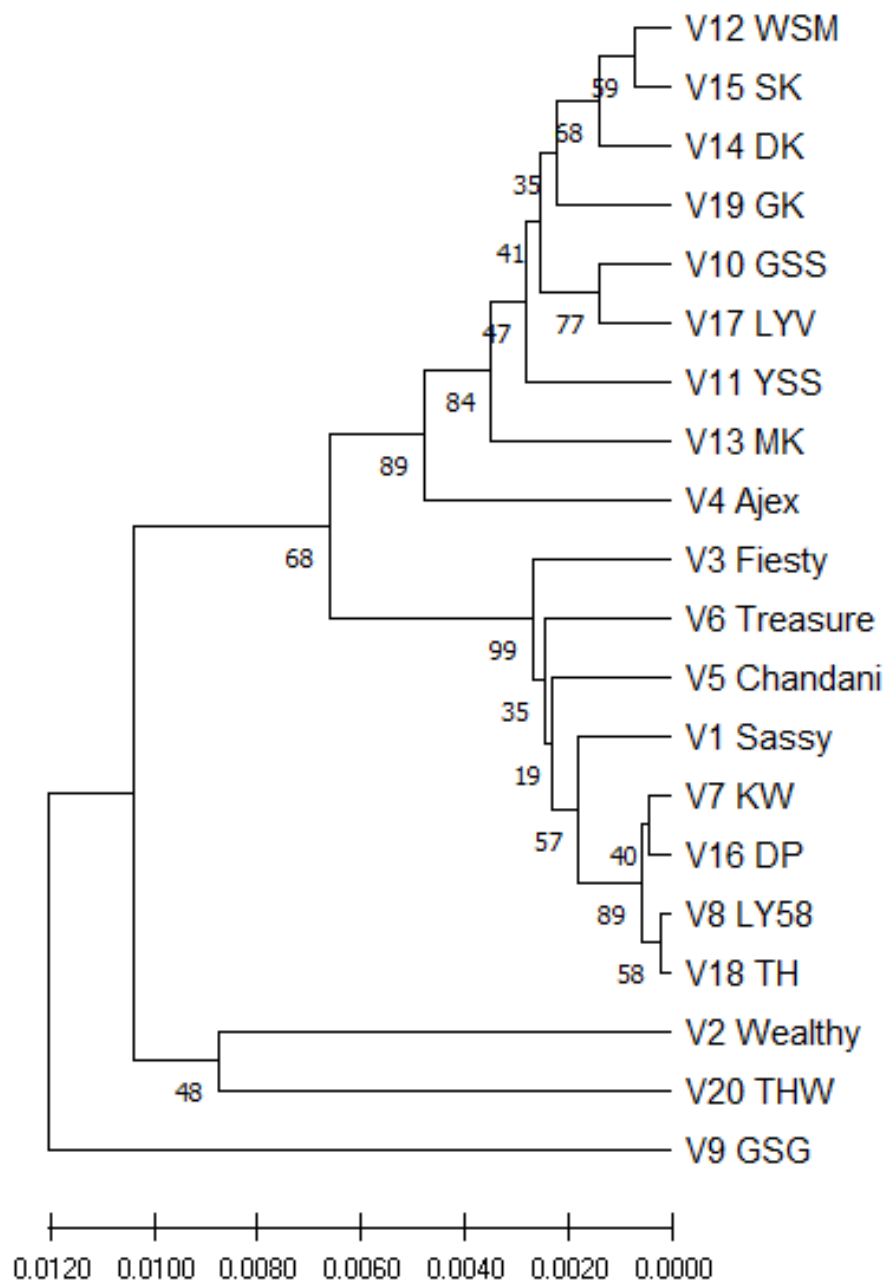


Figure 27: UPGMA tree for *rbcL* and *trnH-psbA* integrated sequences of *Cucumis melo* L. and *Cumis sativus* L.

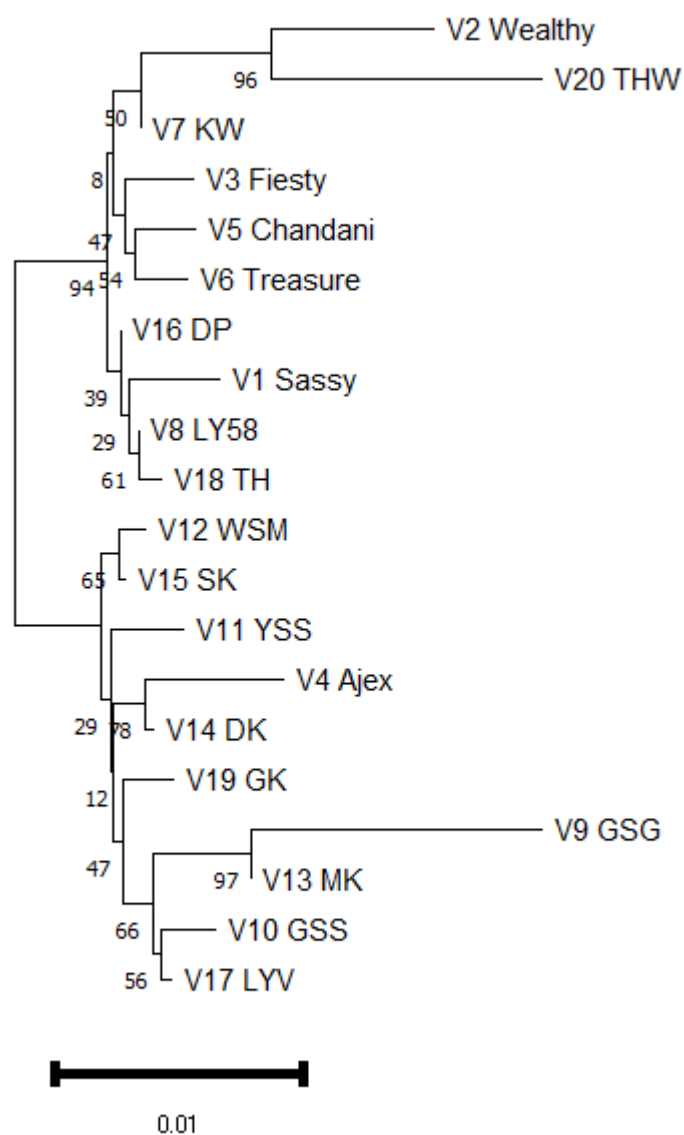


Figure 28: Neighbor joining tree for *rbcL* and *trnH-psbA* integrated sequences of *Cucumis melo* L. and *Cumis sativus* L.

Phylogenetic analysis by including GeneBank data

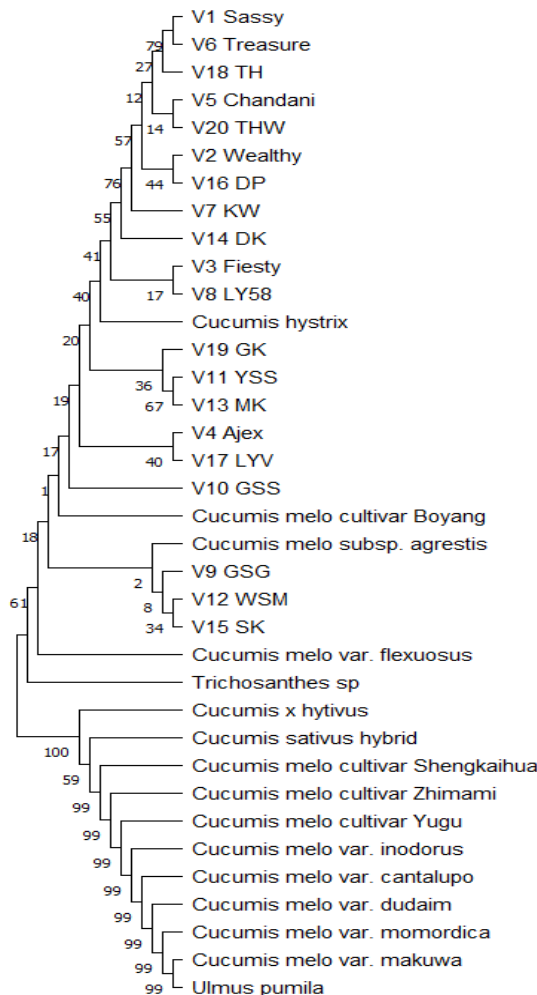


Figure 29: Identification and Comparison of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* L. species with GenBank data, based on integrated *rbcL* and *trnH-psbA* sequences.

In the Neighbor-Joining phylogenetic tree, the genotypes were grouped into distinct clades, with strong bootstrap support (>70) at several nodes. Genotypes *V1* to *V3* and *V5* to *V20* formed separate clusters distinct from *Cucumis melo* L. cultivars. *Cucumis hystrix* L. formed an intermediate clade with *LY58* and *Fiesty*. The phylogenetic analysis reveals clear genetic differentiation between Sri Lankan *Cucumis melo* L. and *Cucumis sativus* L. varieties, as well

as globally cultivated and wild relatives. Further molecular and phenotypic characterization would be valuable to confirm their taxonomic status and evaluate their potential utility in melon breeding programs.

DNA Dating

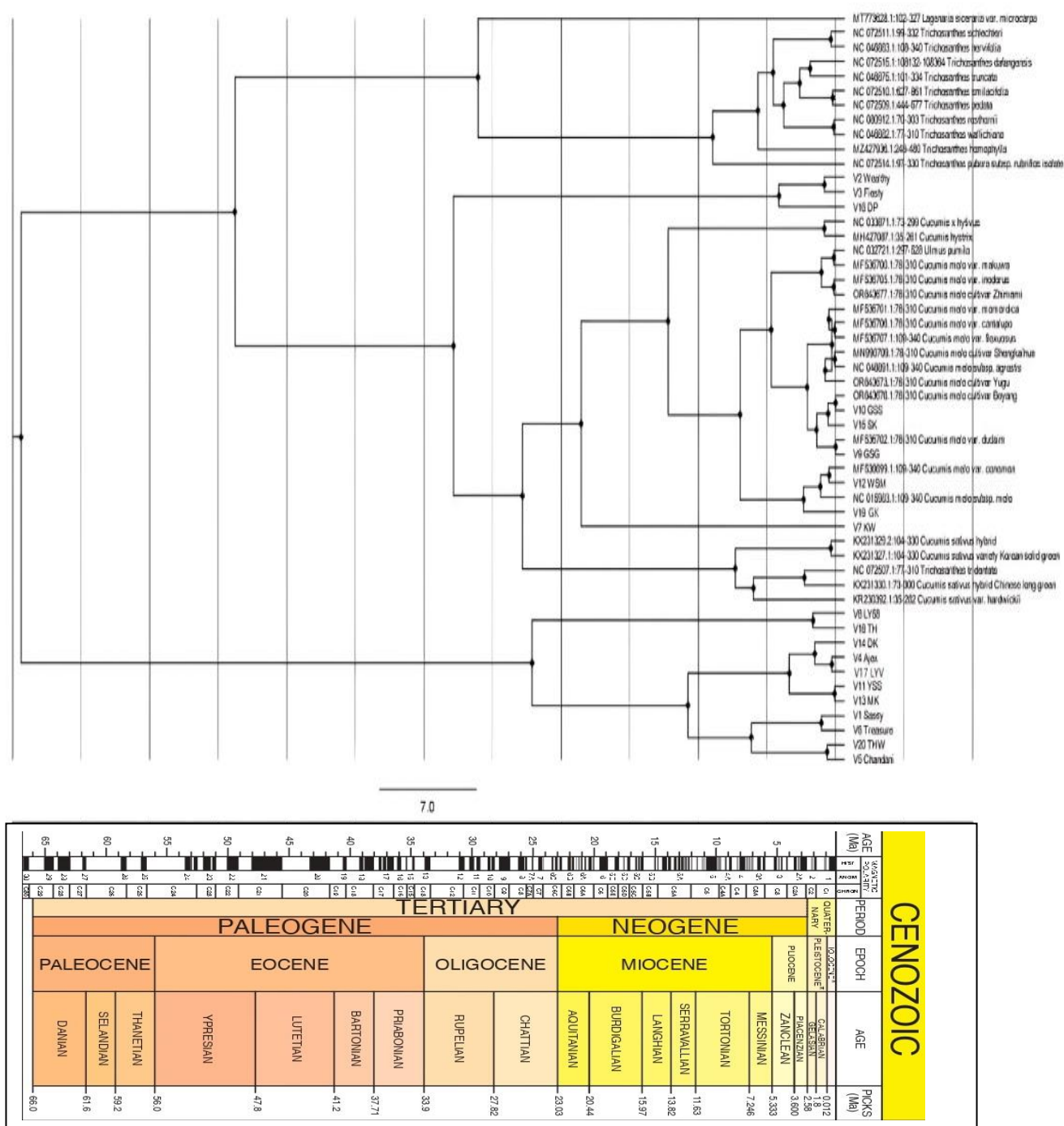


Figure 30: The time calibrated Maximum Clade Credibility tree. The geological time scale is given in the bottom of the diagram. The 95% highest posterior densities node bars are given at the nodes of the tree. The time scale is given below the tree.

In the dated phylogenetic tree, the horizontal axis is labeled with estimated divergence times (in millions of years). The tree shows that varieties such as *Sassy* (V1), *Wealthy* (V2), *Fiesty* (V3), *Ajex* (V4), *Chandani* (V5), *Treasure* (V6), and *KW* (V7) are located on branches near the tips of the tree, with divergence times of less than 3 million years. These varieties have

diverged relatively recently, indicating close evolutionary relationships. They are likely modern cultivars or closely related varieties with minimal genetic differentiation and a recent common ancestor. These varieties have been selected for specific agricultural or phenotypic traits, which align with their morphological similarities.

Varieties such as *LY58* (V8), *GSG* (V9), *GSS* (V10), *YSS* (V11), *WSM* (V12), *MK* (V13), *DK* (V14), *SK* (V15), and *DP* (V16) show divergence times ranging from 3 to 7 million years. This group indicates that these varieties share a more distant common ancestor with the first group but have diverged within the past 7 million years.

According to the dated phylogenetic tree, *LYV* (V17), *TH* (V18), *GK* (V19), and *THW* (V20) are located on branches that diverged between 7 and 20 million years ago. These taxa are more distantly related to the other varieties and separated from their common ancestors much earlier. The greater the divergence time, the more genetic variation is expected. Older-diverging varieties like *LYV*, *TH*, *GK*, and *THW* may possess valuable traits for crop improvement, such as stress tolerance.

Some larger clades in the tree show branching points dating back as far as 60 million years (e.g., the root node). This indicates the age of the family or major groups to which these taxa belong. The leftmost branches represent ancient divergence points, with splits occurring up to 60 million years ago, leading to different genera or major taxonomic groups within the broader *Cucurbitaceae* family, such as *Trichosanthes* and *Lagenaria*. These groups are distantly related to the varieties listed above and have unique evolutionary trajectories, possibly reflecting adaptive radiation or speciation events that occurred tens of millions of years ago. Each group in this tree offers unique genetic resources: closely related varieties are valuable for maintaining yield and quality, while more distantly related species may offer resilience traits beneficial for sustainable agriculture.

Conclusion

This study has provided valuable insights into the genetic, morphological, and sensory characteristics of *Cucumis melo* L. (Vern. Kekiri) and *Cucumis sativus* L. (cucumber) germplasm from Sri Lanka. Nucleotide BLAST analysis revealed that several genotypes, including *Sassy*, *Wealthy*, *Fiesty*, *Ajex*, *Chandani*, *Treasure*, *KW*, *LY58*, *DP*, *TH*, and *THW*, exhibited over 99% identity with *Cucumis sativus*, while *GSG*, *GSS*, *YSS*, *WSM*, *MK*, *DK*, *SK*, and the wild variety *GK* matched *Cucumis melo* with similar identity values. These findings allowed for a clear delineation of species boundaries within Sri Lanka's *Cucumis* germplasm, revealing its genetic diversity on a global scale.

Phylogenetic trees based on *rbcL* and *trnH-psbA* markers identified distinct clades, each with high bootstrap support, providing a robust framework for understanding the evolutionary relationships among the sampled *Cucumis* species. However, some clades showed lower bootstrap support, suggesting the need for additional sequencing data or loci to fully resolve these relationships. The slight variations in tree topology between the two markers emphasize the importance of multi-locus barcoding to gain a more comprehensive understanding of the genetic relationships among *Cucumis* species. These clustering patterns offer valuable insights for selecting genetically diverse parents for breeding programs and support conservation efforts for rare or unique Sri Lankan *Kekiri* accessions.

Our study also revealed the divergence between *Cucumis melo* and *Cucumis sativus*, with findings suggesting that their separation occurred approximately 23 million years ago, contrasting with previous estimates ranging from 4.4 to 11 million years. These discrepancies underscore the complexity of *Cucumis* evolution and highlight the need for further research using additional markers to refine our understanding of these species' evolutionary history.

The statistical analysis of sensory attributes highlights the complexity of consumer preferences in *Cucumis* preparations. Strong correlations were found between color, aroma, and overall performance in cooking varieties, with color and aroma emerging as the most influential sensory traits. For fruit varieties, sweetness was identified as the dominant factor in consumer preference, while texture and color showed weaker correlations. In vegetable salad varieties, a significant negative relationship between bitterness and overall performance was observed, suggesting that increased bitterness reduces consumer satisfaction. This finding offers valuable insights for optimizing pickle or dried cucumber products.

Principal Component Analysis (PCA) demonstrated that consumers evaluate *Cucumis* preparations based on the interplay of multiple sensory attributes rather than isolated traits. Cooking varieties exhibited the most complex relationships between attributes, while fruits emphasized sweetness, and salads balanced flavors like bitterness. The multidimensional nature of sensory evaluation highlights the importance of considering all sensory parameters collectively in product development.

In conclusion, this study underscores the potential of *Cucumis melo* L. as a key crop for agricultural, medicinal, and nutritional applications. The integration of genetic, phenotypic, and sensory data provides a comprehensive framework for future breeding, conservation, and product development initiatives. Continued research into *Cucumis* species, particularly in underutilized regions like Sri Lanka, will be essential for enhancing food security, sustainability, and the development of new varieties with improved traits. The findings of this study contribute significantly to the broader understanding of *Cucumis* genetic diversity and lay a strong foundation for future research and development in both agriculture and medicine.

Acknowledgments

Genetech Sri Lanka for Molecular Laboratory work. Mr. Roshan Weerakoon, Mrs. Manjula Godakandha and Base hospital Cheddikulam, The medical superintendant Dr. Akilendran and minor staff for facilitating field trials.

Supporting information

Table 1: List of collected genotypes (11 *C. melo* genotypes and 9 *C. sativus* genotypes) and their acronyms.

Table 2: Ingredients of prepared *Cucumis melo* L. and *Cucumis sativus* L. dishes

Table 3: Ingredients of prepared salad dishes

Table 4: Evaluated key plant DNA barcoding markers.

Figure 1: Field planting layout. 19 Genotypes (G), two replicates (B or R), and four plants (P) per replicate are shown here. All (11 *C. melo* and eight *C. sativus*) genotypes were assessed together.

Figure 2: Morphological variation of the reproductive parts (A: leaves; B: flowers; C: fruits appearance; D: cross sections of fruits; E: seeds). The numbers 1–14 indicate Sassy; Wealthy; Fiesty; Ajex; Chandani; Treasure; KW; LY58; GSG; GSS; YSS; WSM; MK; DK.

Figure 3: Morphological variation of the reproductive parts (A: leaves; B: flowers; C: fruits appearance; D: cross sections of fruits; E: seeds). The numbers 15–10 indicate SK; DP; LYV; TH; GK; THW.

Figure 4: Comparative growth rates of different varieties, showing variability in performance across genotypes. Highlights the superior growth of *WSM* (V12), *SK* (V15), and *YSS* (V11) compared to *Sassy* (V1), *MK* (V13), and *DP* (V16).

Figure 5: Boxplot analysis of morphological parameters in early developmental stages. Demonstrates the variability in SGP (%), NDSV, NDSFl, and NDSFr

Figure 6: Boxplot analysis of morphological parameters in early developmental stages. Demonstrates the variability in LV, NN, VV, NL, SLB, LM, HS, ODR, and OPR.

Figure7: Plate: Culinary preparations subjected to taste panel and subsequent association analyses between sensory parameters and six cooking *Cucumis* varieties.

Figure 8: Dish: Fruit salad preparations subjected to taste panel and subsequent association analyses between sensory parameters and *Cucumis melo* L. fruit varieties.

Figure 9: Scree Plot with Cumulative Variance. Scree plot depicting the explained variance for each principal component in the organoleptic assessment analysis. The blue bars represent the percentage of variance explained by individual principal components, while the red line shows the cumulative explained variance. This visualization helps determine the optimal number of principal components to retain for interpreting the sensory attribute data, with components accounting for substantial cumulative variance providing the most meaningful insights into the underlying data structure.

Figure 10: PCA Biplot of Sensory Attributes. Principal Component Analysis (PCA) biplot revealing the relationships between sensory attributes: Color, Aroma, Texture, Bitterness, and Overall Performance. The blue scatter points represent individual data points projected onto the first two principal components (PC1 and PC2). Red arrows indicate the direction and magnitude of each sensory attribute's contribution to the principal components. The length and orientation of these vectors suggest the relative importance and correlation between attributes, with closer-aligned vectors indicating stronger relationships. The axes display the percentage of variance explained by PC1 and PC2, providing context for the dimensionality reduction.

Figure 11: Scree plot illustrating the explained variance for each principal component in the fruit organoleptic assessment. The bar graph displays the percentage of variance explained by individual principal components, providing insight into the dimensionality reduction process. This visualization helps researchers identify the most significant components that capture the underlying variability in sensory attributes, crucial for understanding the data's structure and complexity.

Figure 12: PCA Biplot. Principal Component Analysis (PCA) biplot revealing the multivariate relationships between key sensory attributes: Color, Aroma, Texture, Sweetness, and Overall Performance. The scatter plot represents individual data points projected onto the first two principal components (PC1 and PC2). Red arrows indicate the direction and magnitude of each attribute's contribution to the principal components, with arrow length and orientation suggesting the relative importance and potential correlations between different sensory characteristics.

Figure 13 :Correlation Matrix Heatmap. Correlation matrix heatmap displaying the inter-relationships between sensory attributes using a color-coded scale. The visualization uses a diverging color palette (cool to warm) to represent the strength and direction of correlations between Color, Aroma, Texture, Sweetness, and Overall Performance. Numerical annotations provide precise correlation coefficients, allowing for a detailed interpretation of how different sensory attributes relate to one another in the fruit organoleptic assessment.

Figure 14: PCA Scatter Plot. Principal Component Analysis (PCA) scatter plot revealing the distribution of different salad variants across the first two principal components. Each point represents a salad sample, color-coded by salad type to illustrate the sensory attribute variations. The plot demonstrates the multivariate differences between salad samples, showing how they cluster or diverge based on their organoleptic characteristics.

Figure 15: Scree Plot. Scree plot illustrating the proportion of variance explained by each principal component in the salad organoleptic assessment. The line graph displays the cumulative explained variance, helping identify the most significant components that capture the underlying variability in sensory attributes.

Figure 16: Correlation Biplot. Correlation biplot depicting the relationships between sensory attributes and the first two principal components. Blue scattered points represent individual salad samples, while red arrows indicate the direction and magnitude of each sensory attribute's contribution to the principal components. The text labels show the attributes' positions, revealing how different sensory characteristics are interrelated and contribute to the overall variance in the dataset.

Figure 17: Visualization of the *rbcL* PCR products for *Cucumis* varieties DNA, obtained from fresh leave samples. 1: molecular marker (100 bp ladder), 2: Sassy, 3: Wealthy, 4: Fiesty, 5: Ajex, 6: Chandani, 7: Treasure, 8: KW, 9: LY58, 10: GSG, 11: GSS, 12: YSS, 13: WSM, 14: MK, 15: DK, 16: SK, 17: DP, 18: LYV, 19: TH, 20: molecular marker (100 bp ladder), 21: molecular marker (100 bp ladder), 22: GK, 23: THW, 24: Distilled water.

Figure 18: Visualization of the *psbA-trnH* PCR products for *Cucumis* varieties DNA, obtained from fresh leave samples. 1: molecular marker (100 bp ladder), 2: Sassy, 3: Wealthy, 4: Fiesty, 5: Ajex, 6: Chandani, 7: Treasure, 8: KW, 9: LY58, 10: GSG, 11: molecular marker (100 bp ladder), 12: molecular marker (100 bp ladder), 13: GSS, 14: YSS, 15: WSM, 16: MK, 17: DK, 18: SK, 19: molecular marker (100 bp ladder), 20: DP, 21: LYV, 22: TH, 23: GK, 24: THW, 25: Distilled water PCR.

Figure 19: DNA sequencing of *rbcL* and *trnH-psbA* pop sets of *Cucumis melo* L. and *Cucumis sativus* L.

Figure 20: DNA sequence alignment diagram for *rbcL* region of the chloroplast genome of *Cucumis melo* L. and *Cucumis sativus* L. cultivars. The alignment was obtained by using MEGA 11 software. The DNA sequence differences shown in this illustration were used to draw the genetic dissimilarity diagram (i.e. dendrogram) shown in Diagram 1.

Figure 21: DNA sequence alignment diagram for *trnH-psbA* region of the chloroplast genome of *Cucumis melo* L. and *Cucumis sativus* L. cultivars. The alignment was obtained by using MEGA 11 software. The DNA sequence differences shown in this illustration were used to draw the genetic dissimilarity diagram (i.e. dendrogram) shown in Diagram 2.

Figure 22 : The genetic dissimilarity diagram (i.e. dendrogram) for *rbcL* region of the chloroplast genome of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* cultivars L.

Figure 23: The genetic dissimilarity diagram (i.e. dendrogram) for *trnH-psbA* region of the chloroplast genome of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* cultivars L.

Figure 24: Base locations of SNPs, INDELs, and shared bases are highlighted for easy visualization for *rbcL* and *trnH-psbA* locus in *Cucumis melo* L. and *Cucumis sativus* germplasms in Sri Lanka.

Figure 25: Molecular Phylogenetic affinity analysis of *rbcL* gene sequences of *Cucumis melo* L. and *Cucumis sativus* L. plants of Sri Lanka. The evolutionary history inferred using the Neighbor-Joining method. The tree drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances computed using the Maximum composite likelihood method and are in the units of the number of base substitutions per site. The analysis included 20 nucleotide sequences of 20 plants and evolutionary analyses were conducted in MEGA11.

Figure 26: Molecular Phylogenetic affinity analysis of *trnH-psbA* gene sequences of *Cucumis melo* L. and *Cucumis sativus* L. plants of Sri Lanka. The evolutionary history inferred using the Neighbor-Joining method. The tree drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances computed using the Maximum composite likelihood method and are in the units of the number of base substitutions per site. The analysis included 20 nucleotide sequences of 20 plants and evolutionary analyses were conducted in MEGA11.

Figure 27: UPGMA tree for *rbcL* and *trnH-psbA* integrated sequences of *Cucumis melo* L. and *Cucumis sativus* L.

Figure 28: Neighbor joining tree for *rbcL* and *trnH-psbA* integrated sequences of *Cucumis melo* L. and *Cucumis sativus* L.

Figure 29: Identification and Comparison of Sri Lankan *Cucumis melo* L. and *Cucumis sativus* L. species with GenBank data, based on integrated *rbcL* and *trnH-psbA* sequences.

Figure 30: The time calibrated Maximum Clade Credibility tree. The geological time scale is given in the bottom of the diagram. The 95% highest posterior densities node bars are given at the nodes of the tree. The time scale is given below the tree.

References

- Edgar, R.C. (2004). MUSCLE: Multiple Sequence Alignment with High Accuracy and High Throughput. *Nucleic Acids Research*, 32, 1792-1797. <https://doi.org/10.1093/nar/gkh340>
- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33, 1870-1874. <https://doi.org/10.1093/molbev/msw054>
- Kimura, M. (1980). A Simple Method for Estimating Evolutionary Rates of Base Substitutions through Comparative Studies of Nucleotide Sequences. *Journal of Molecular Evolution*, 16, 111-120. <https://doi.org/10.1007/BF01731581>
- Cognato, A.I. (2013). Molecular Phylogeny and Taxonomic Review of *Premnobiini* Browne, 1962 (Coleoptera: Curculionidae: Scolytinae). *Frontiers in Ecology and Evolution*, 1: 1.
- Dassanayake, M.D., and Fosberg, F.R. (1988). *A Revised Handbook of the Flora of Ceylon*. CRC Press, Boca Raton.
- Diez, M.J., van Dooijeweert, W., Maggioni, L., and Lipman, E. (2005). Report of a Working Group on Cucurbits. *European Cooperative Programme for Plant Genetic Resources (ECPGR)*. Available at: https://www.bioversityinternational.org/fileadmin/_migrated/uploads/tx_news/Report_of_a_Working_Group_on_Cucurbits_1282.pdf. Accessed on July 25, 2021.
- Drummond, A.J., and Rambaut, A. (2007). BEAST: Bayesian Evolutionary Analysis by Sampling Trees. *BMC Evolutionary Biology*, 7: 214. <https://doi.org/10.1186/1471-2148-7-214>
- Hollingsworth, P.M., Graham, S.W., and Little, D.P. (2011). Choosing and Using a Plant DNA Barcode. *PloS ONE*, 6(5).
- Huelsenbeck, J.P., and Ronquist, F. (2001). MrBayes: Bayesian Inference of Phylogeny. *Bioinformatics*, 17, 754-755.
- Knerr, L.D., Staub, J.E., Holder, D.J., and May, B.P. (1989). Genetic Diversity in *Cucumis sativus* L. Assessed by Variation at 18 Allozyme Coding Loci. *Theoretical and Applied Genetics*, 78(1): 119-128.
- Kocyan, A., Zhang, L.B., Schaefer, H., and Renner, S.S. (2007). A Multi-Locus Chloroplast Phylogeny for the Cucurbitaceae and Its Implications for Character Evolution and Classification. *Molecular Phylogenetics and Evolution*, 44(2): 553-577.
- Kramer, D.M., Johnson, G., Kiirats, O., and Edwards, G.E. (2004). New Fluorescence Parameters for the Determination of *Q A* Redox State and Excitation Energy Fluxes. *Photosynthesis Research*, 79(2): 209-218.
- Kress, W.J., Erickson, D.L., Jones, F.A., Swenson, N.G., and Perez, R. (2009). Plant DNA Barcodes and a Community Phylogeny of a Tropical Forest Dynamics Plot in Panama. *Proceedings of the National Academy of Sciences*, 106: 18621–18626. <https://doi.org/10.1073/pnas.0909820106>
- Kuhlert, S., Austic, G., Zegarac, R., Osei-Bonsu, I., Hoh, D., Chilvers, M.I., Roth, M.G., Bi, K., TerAvest, D., Weebadde, P., and Kramer, D.M. (2016). MultispeQ Beta: A Tool for Large-Scale Plant Phenotyping Connected to the Open PhotosynQ Network. *Royal Society Open Science*, 3(10): p.160592.

- Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, 33(7): 1870–1874.
- Kumari, S.A.S.M., Nakandala, N.D.U.S., Nawanjana, P.W.I., Rathnayake, R.M.S.K., Senavirathna, H.M.T.N., Senevirathna, R.W.K.M., Wijesundara, W.M.D.A., Ranaweera, L.T., Mannanayake, M.A.D.K., Weebadde, C.K., and Sooriyapathirana, S.D.S.S. (2019). The Establishment of the Species-Delimits and Varietal-Identities of the Cultivated Germplasm of *Luffa acutangula* and *Luffa aegyptiaca* in Sri Lanka Using Morphometric, Organoleptic and Phylogenetic Approaches. *PloS ONE*, 14(4).
- Lester, G. (1996). Melon (*Cucumis melo* L.) Fruit Nutritional Quality and Health Functionality. *Hortscience*, 31(4): 693.
- Levin, R.A., Wagner, W.L., and Hoch, P.C. (1992). Family-Level Relationships of *Onagraceae* Based on Chloroplast *rbcL* and *ndhf* Data. *American Journal of Botany*, 90: 107-115. <https://doi.org/10.3732/ajb.90.1.107>
- Li, L., Madriñán, S., and Li, J. (2016). Phylogeny and Biogeography of *Caryodaphnopsis* (Lauraceae) Inferred from Low-Copy Nuclear Gene and ITS Sequences. *Taxon*, 65(3): 433-443.
- Liu, L., Kakihara, F., and Kato, M. (2004). Characterization of Six Varieties of *Cucumis melo* L. Based on Morphological and Physiological Characters, Including Shelf-Life of Fruit. *Euphytica*, 135(3): 305.
- Lv, J., Qi, J., Shi, Q., Shen, D., Zhang, S., Shao, G., Li, H., Sun, Z., Weng, Y., Shang, Y., and Gu, X. (2012). Genetic Diversity and Population Structure of Cucumber (*Cucumis sativus* L.). *PloS ONE*, 7(10).
- Maggs, G.K., Madsen, S., and Christiansen, J.L. (2000). Genetic Marker Techniques in the Family *Cucurbitaceae*. *Genetic Resource and Crop Evolution*, 47: 385-393.
- Miller, M.A., Pfeiffer, W., and Schwartz, T. (2010). Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees. *Gateway Computing Environments Workshop (GCE)*. Available from: http://www.ieeexplore.ieee.org/abstract_document/5676129.
- Mliki, A., Staub, J.E., Zhangyong, S., and Ghorbel, A. (2001). Genetic Diversity in Melon (*Cucumis melo* L.): An Evaluation of African Germplasm. *Genetic Resources and Crop Evolution*, 48(6): 587-597.
- Nakata, E., Staub, J.E., López-Sesé, A.I., and Katzir, N. (2005). Genetic Diversity of Japanese Melon Cultivars (*Cucumis melo* L.) as Assessed by Random Amplified Polymorphic DNA and Simple Sequence Repeat Markers. *Genetic Resources and Crop Evolution*, 52(4): 405-419.
- Okubo, H., Oohama, K., Inokuchi, S.I., and Fujieda, K. (1990). *Kekiri*, an Introduction of Cooking Melon from Sri Lanka and Its Taxonomic Rank in the Genus *Cucumis*. *Scientia Horticulturae*, 43(3-4): 273-280.
- Posada, D. (2008). JModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution*, 25: 1253-1256.
- Rajapaksha, U. (1998). *Traditional Food Plants in Sri Lanka*. Colombo, Sri Lanka: Hector Kobbekaduwa Agrarian Research and Training Institute.
- Rambaut, A. (2014). *FigTree*, a Graphical Viewer of Phylogenetic Trees. Available from: <http://tree.bio.ed.ac.uk/software/figtree>.
- Stamatakis, A. (2006). RAxML-VI-HP: Maximum Likelihood-Based Phylogenetic Analyses with Thousands of Taxa and Mixed Models. *Bioinformatics*, 22(21): 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>

***FUNDING DECLARATION: NO FUNDING WAS RECEIVED FOR THIS RESEARCH.**

***AN AI TOOL (Chat GPT) WAS USED TO CORRECT ENGLISH AND REFINING.**